

Another piece of paper without force

This week's Start II treaty signed in Moscow is little more than wishful thinking, at least while it lacks the support of other ex-Soviet republics than Russia.

IN years past, an agreement between the United States and Russia to dispense with two-thirds of existing nuclear weapons would be hailed as the beginning of a new era of peace and tranquillity. But that reading cannot be wrung from the signing ceremony last Sunday in Moscow of the treaty called Start II. President Boris Yeltsin of Russia, bruised by last month's congress, would have welcomed the chance to appear, yet again, as a personage of international importance, while President George Bush of the United States was no doubt alert to the history books that will engulf him on Inauguration Day, 20 January.

Start II is a splendid piece of paper, but it will be a dead letter until ratified not only by the Russian Congress, but by the other *de facto* nuclear weapons states of the Commonwealth of Independent States (CIS), the constitutional legatee of what was once the Soviet Union. The trouble is that the CIS seems never to meet. Even last year's Start I treaty has not yet won its approval. The US Senate is hardly likely to give the treaties a hearing until there is some action in Minsk (the formal seat of the CIS). Peace and tranquillity are postponed until then.

So the true significance of this week's signing ceremony is the reminder it should provide of the ambiguity of the ex-Soviet republics on nuclear weapons. During the past year, the republics have been zealous at appropriating to themselves ex-Soviet equipment lying on their territory. The Ukraine, in military strength next only to Russia, has said it has no wish to be a free-standing nuclear power, but has asserted rights to the weapons left on its territory. As time passes and nothing much happens at Minsk, the Ukraine and other republic governments must be tempted to regard their "own" nuclear weapons as a symbolic counterpoise to the power that remains in Moscow. At the very least, Yeltsin would have to ask them nicely even formally to agree to a continuation of present arrangements for control and command of them. The republics would not readily give up the right to ask for something in return.

That is why the most urgent task for the diplomats who have worked up Start II is to create the kind of forum in which the republic governments can be properly engaged in discussions of the strategic relationship of the CIS with the outside world. If Minsk remains inactive, why not arrange a separate forum, at Geneva or elsewhere? Boringly for those concerned, in such a setting it would be necessary to re-educate a new set of participants from the East in the arguments for believing that Start II is in everybody's

interests. Meanwhile, Sunday's signing ceremony will have been symbolic only. □

Gallo on the rack

The latest government pronouncement raises questions about the office itself as well as about Gallo.

THE dictum that to be just, justice must be swift seems not to apply to Robert C. Gallo, the researcher at the US National Institutes of Health (NIH). The latest opinion (see page 3) from the Office of Research Integrity (ORI) within NIH's parent body, the Department of Health and Human Services, is another setback for him, but an inconclusive one. Gallo says that the half-sentence now called a "lie" has been misunderstood. In due course, his lawyers may help him make his explanation stick. Last May, a committee of the NIH accepted it.

Last May's report, leaked in unfinished form, was most damaging of Gallo for what it had to say about the management of his laboratory during the period in the 1980s when the hunt for the AIDS virus engaged everybody's attention there. It was a laboratory in which mistakes such as the misattribution of samples could easily have happened. That does not, of course, imply that Gallo or anybody else knowingly misappropriated Montagnier's virus, then called LAV. Nor does sloppiness constitute misconduct. But that also explains why Gallo has been unable to prove that his virus and that of Luc Montagnier of the Institut Pasteur have independent origins. Gallo's initial disbelief of Montagnier's claim to have isolated a virus from AIDS patients, which he has since acknowledged to have been unfortunate, inevitably colours the dispute that has rumbled on since 1984.

What should happen now? Gallo's energy and resources were crucial in the speedy identification of the virus responsible for AIDS, and for the diagnostic tools in use since the second half of 1984. But Gallo's estimation of his own laboratory's progress was immodest, while Montagnier deserved more credit than Gallo gave him. It is futile to expect that further parsing of the literature can decide whether Gallo and/or his associates knowingly used Montagnier's virus as their own. The practical question now is simply whether the equal division of royalties between France and the United States, agreed in 1986, should persist. The notion that a federal investigative office can function as a universal Solomon is even less tenable now than when it belonged to NIH. □



John Gibbons, US science adviser, gets high marks for political skills

Washington. John Gibbons, named last week as science adviser to President-elect Bill Clinton and director-designate of the White House Office of Science and Technology Policy, fits the dominant profile of those joining the new administration: an experienced and respected administrator, a team player and someone with strong ties to Congress. And while scientists are pleasantly surprised that he was selected as part of the initial round of cabinet appointments, their happiness is tempered by uncertainty about the fate of research under the new administration.

Without exception, those who know the 63-year-old Gibbons and his 13 years as head of the Office of Technology Assessment (OTA) hold both in high regard. "It's a choice made in heaven", says Jeremy Stone, president of the Federation of American Scientists, one of three scientific organizations to have honoured Gibbons recently for outstanding public service (in 1991 the American Physical Society (APS) gave him its Leo Szilard Award, and last year the American Association for the Advancement of Science bestowed its Philip Abelson Prize on him). "He's a perfect match for a president and a vice president who want to understand complex issues, and he's survived the political wars that go along with some of these issues."

Gibbons is a former nuclear physicist turned public servant who is in the midst of his third six-year term as director of OTA, an arm of the US Congress that churns out one major report a week on topics ranging from genetic screening to nuclear weapons. Gibbons is widely credited with rescuing the 20-year-old agency from the political squabbling that crippled it during its early years and turning it into an effective, non-partisan and parsimonious government agency. His early career was spent at Oak Ridge National Laboratory, studying stellar evolution and the origins of heavy elements in the Solar System. During the 1970s he worked on energy and environmental policy, in particular on conservation and the efficient use of energy.

With Vice President-elect Al Gore expected to be the chief architect of administration policy on science and technology, Gibbons is seen as an excellent choice for a supporting role. "He'll be presenting the options to Al Gore, just as at OTA he presented options to Congress", says APS's Bob Park. Adds Robert Fri, president of

Resources for the Future, an economic think tank in Washington which Gibbons served as a member of the board of directors for nearly a decade, "he's not bland, and he's perfectly capable of speaking his mind. But he's not someone who needs to see his name

technology as a way of strengthening the US economy and improving the country's standard of living. What the new administration has not said, however, is whether that support will come at the expense of basic research.

Last year, Gibbons said that the ultimate rationale for federal spending on basic research is its relevance to broad, national goals. But Gibbons also criticized such expensive scientific projects as the Superconducting Super Collider and the Space Station Freedom — projects with important political constituencies that Clinton and Gore support — for "starving" individual investigators in those fields, and he said that reports calling for increased federal spending on research "are simply whistling in the shadow" of the \$350-billion-a-year federal deficit.

At OTA, Gibbons was adept at combining disparate views and producing reports seen as comprehensive, balanced and useful to policy makers. As science adviser his job will be to serve the new president in a similar manner at the same time he tries to make Clinton's decisions palatable to the scientific community.

Jeffrey Mervis



Gibbons is at home with data.

and his picture in the newspaper every week."

Although Gibbons has declined to grant interviews since his appointment, he is believed to share Clinton's view that the government should support advanced

Healy: In or out at NIH?

Washington. One of the more entertaining games for biomedical researchers these days is speculating on whether US President-elect Bill Clinton will remove Bernadine Healy as director of the National Institutes of Health (NIH). The heavy betting at this stage is that he will.

In the past few weeks the Clinton transition team has asked a number of prominent research physicians if they would be interested in Healy's job, adding to previous rumours that Clinton plans to replace her. Last week, however, Donna Shalala, his nominee to head the Health and Human Services Department (HHS) of which NIH is a part, told the transition team to stop its inquiries. It is not clear whether that message was an implicit endorsement of the incumbent or merely a sign from Shalala, a member of the NIH director's advisory board, that things were moving too quickly.

No one had reached the top of the list before Shalala's intervention, mainly because the NIH position is not yet a priority for the new administration. And some

potential candidates have already said "no thank you". The idea that Healy would be safe because she is a woman has faded in the face of Clinton's appointment of two women (Shalala and Surgeon General-designate Jocelyn Elders) to higher-ranking positions within HHS.

Healy, who joined NIH in April 1991 after her appointment by Republican President George Bush, has not won wide support. Although she took the politically correct route of establishing a multi-million-dollar "women's health initiative", she has failed to win the enduring support of the women's political caucus in Congress. Her conscience-driven but politically risky confrontations with US Representative John Dingell (Democrat-Michigan) has put him on her list of enemies. And her "strategic plan" for biomedical research was seen as an attempt to impose NIH as 'Big Brother', weakening her support within a scientific community that, in previous years, has lobbied to keep the NIH director in place during presidential transitions.

Barbara J. Culliton

Japanese science stands out as recession squeezes budget

Tokyo. Japan's science-related ministries and agencies have managed to receive comparatively healthy budgets for 1993 despite a clampdown on government spending because of the recession. But some of the most ambitious proposals, in particular from the

Sunshine Project, which covers research on alternative energy, energy conservation and 'environment friendly' technology. The Industrial Scientific Technology programme, covering the ministry's 'large-scale' (*ogata*) and 'next-generation' (*jiseidai*) projects and attracting foreign as well as Japanese companies (see *Nature* **360**, 500; 1992), received an increase of 7.2 per cent. But at the same time, the Human Frontier Science Program, which is funded jointly with the Science and Technology Agency (STA) and which supports international research on the brain and molecular biology, was held practically constant, at ¥3.9 billion.

The Ministry of Education, Science and Culture (MESC) received most of its request for a 16 per cent increase in the budget for research grants for university researchers. Its 14 per cent increase, the largest in more than a decade, will raise its budget to ¥73.6 billion.

The Ministry of Health and Welfare (MHW) also succeeded in getting a significant increase in its budget to fight AIDS through research, screening, counselling, local government subsidies and public education. But even with a fivefold increase,

to just over ¥10 billion (US\$80 million), its budget is tiny compared with those in the West.

STA also will receive 20 per cent more for its 'Special Promotion Funds' to provide grants for government researchers. But its budget of ¥13.3 billion is still small compared with that of MESC.

The agency's space budget continues to rise rapidly because of Japan's commitment to the US Space Station Freedom while nuclear energy, which consumes over half of the STA's budget, has been given an increase of only 2.7 per cent. However, within this outlay the budget for the International Thermonuclear Experimental Reactor (ITER) project will increase by nearly 30 per cent as Japan begins the engineering design phase of ITER at Naka in Ibaraki Prefecture (see *Nature* **360**, 615; 1992). The agency's small (¥1.6 billion) budget for the human genome project will increase by 45 per cent, with most of the extra money going towards a computer link with the database at John Hopkins University in Baltimore.

David Swinbanks

US report finds NIH's Gallo guilty of misstatement

Washington. In a new twist to the case of Robert C. Gallo and the AIDS virus, a high-level review body of the US Department of Health and Human Services (HHS) has partially overturned an earlier report that vindicated the National Institutes of Health (NIH) researcher of scientific misconduct in a long-running dispute with scientists from the Pasteur Institute in Paris. Specifically, the HHS's Office of Research Integrity (ORI) says in a report issued last week that Gallo lied when he wrote in a 1984 paper that a putative AIDS virus (then called LAV) developed by Luc Montagnier of the Pasteur had not been transmitted to a permanent cell line.

The dispute centres on the fact that LAV and HTLV-IIIb are structurally nearly identical, even though in the early stages of AIDS research they seemed to be biologically different — one grew in culture and one did not. The confusion was cleared up more than a year ago when unpublished data from Montagnier's laboratory revealed that both research groups had actually been working with a virus designated LAI that contaminated the LAV sample at Pasteur before it was sent to Gallo.

In May 1992, NIH cleared Gallo of charges of scientific misconduct for allegedly "misappropriating" an AIDS virus from French collaborators. But it found his colleague, Mikulas Popovic, guilty of misconduct on minor counts regarding the way in which data were reported in a paper published in May 1984 in *Science* (see *Nature* **357**, 3; 1992).

These conclusions came from NIH's Office of Scientific Integrity and were approved by NIH director Bernadine Healy, who forwarded the NIH report to the newly formed Office of Research Integrity in NIH's parent Department of Health and Human Services. Reviewing the case herself, Healy found no evidence of fraud and said so. But her position plainly angered US Representative John Dingell (Democrat-Michigan), who accused Healy of conducting a whitewash and who called the NIH report a "deeply flawed" document.

This posed a dilemma for ORI. If it rejected Healy's endorsement of the report, ORI would be criticizing her judgement and that of other NIH officials who also stood behind Gallo. But if it supported Healy, it would have to explain itself to Dingell.

The new report finds Gallo guilty of one count of scientific misconduct for allegedly falsifying part (but not all) of a sentence in the *Science* paper. Gallo wrote that the apparent differences between his virus (then called HTLV-IIIb) and Montagnier's might be "due to insufficient characterization of

What Japan will spend on science in 1993

(in billion yen; 125 yen = US\$1)

MITI	Amount	Change
Overall R&D	281.9	+ 8.7%
Industrial technology	25.3	+ 7.2
New Sunshine Project	53.9	+ 7.2
Human Frontier Program	3.9 *	>+ 1.0
* shared with Science and Technology Agency		
STA		
Overall R&D	581.5	+ 5.4%
Special Promotion Funds	13.3	+ 20.9
Space	157.0	+ 8.5
Nuclear Energy	323.6	+ 2.7
ITER	6.9	+ 27.8
Ocean Research	12.0	+ 7.1
ERATO	6.9	+ 9.5
Human Genome	1.6	+ 45.5
MESC		
Grants-in-aid of research	73.6	+ 13.9
AIDS	10.1	+ 480.0

Ministry of International Trade and Industry (MITI), have been quite severely trimmed, and the date for achieving the goal of doubling the government budget for science and technology is receding far into the future.

The total budget for fiscal year 1993, which starts on 1 April, was set last week by the cabinet at just over ¥72,000 billion (US\$580 billion), an increase of only 0.2 per cent from 1992 and the smallest in six years. Proposals for science and technology were treated favourably, however, in line with a commitment last year to double the government budget for science and technology "as soon as possible". The budget must be approved by the Diet in the next few months, but funding for science and technology is unlikely to change.

Despite the favoured status of science, MITI received only slightly more than half (8.7 per cent) of its request for an increase of 16 per cent in research and development, to more than ¥300 billion (US\$2.4 billion). Nevertheless, MITI has succeeded in giving a large boost to its New

LAV because the virus has not been transmitted to a permanently growing cell line for true isolation...".

Popovic, who in 1983 and 1984 was the first anywhere to get an AIDS virus to grow in sufficient quantity to develop a test to detect the virus in the blood supply, had actually transmitted LAV to a cell line, although it did not grow well or for long and the Gallo group did not attempt to characterize it. Gallo says the sentence is intended to refer only to the fact that the French had not grown LAV in a permanent cell line (which they had not). The May NIH/OSI report decided that the reference was ambiguous but found no evidence to refute Gallo's explanation.

But ORI concluded instead that "Gallo falsely reported the status of LAV research" in the paper and is therefore guilty of scientific misconduct. In an odd twist, the ORI report simultaneously concurs with the NIH/OSI finding that Popovic is guilty while also coming very close to saying it does not matter. "The confirmed scientific misconduct on the part of Dr Popovic is relatively minor", it states, "does not invalidate the findings of his breakthrough research and should not preclude his employment as a scientist." (Popovic has been out of work for nearly two years.)

Gallo, who says that he will appeal, calls ORI's interpretation of the controversial sentence "utterly unwarranted" and the investigators "incompetent". Popovic is also expected to appeal. "There is no evidentiary basis for any finding of scientific misconduct", says his attorney, Barbara F. Mishkin.

And so the case goes on, with energy on both sides of the Atlantic being diverted from the fight against AIDS.

Barbara J. Culliton

EC increases Framework budget

London. Research ministers from the 12 member states of the European Communities (EC), under pressure to reach a decision by the end of the year, agreed two days before Christmas to spend more on the final two years of the Third Framework Programme, through which the EC Commission supports its joint research projects.

The ministers agreed to provide an extra ECU900 million (US\$1.2 billion), spread over 1993 and 1994. The new money will mean an increase of 13.3 per cent for most of the 15 programmes under the Framework umbrella. However, fusion energy (which is on a different budget cycle) will receive an additional 24 per cent and non-nuclear energies will get 38 per cent more.

The final figure is a compromise between an increase of ECU750 million that the British government wanted and as much as ECU1.5 billion sought by the European Commission.

David Dickson

Forecast 1993

Nature takes a look into its crystal ball at prospects over the next 12 months in several important areas relating to research and the scientific community.

Eastern Europe

Three years after the fall of the German Democratic Republic triggered a domino collapse of communist states in Central and Eastern Europe, celebrations of the new year have lost their sparkle. The arrival of

struggling to establish a science base in their new democracies.

Money is in short supply everywhere, but science reform and restructuring (usually a euphemism for redundancies) have proceeded very gradually because of social resistance. Most Central and Eastern European countries had followed the communist model, itself based on the French model, of separating research from higher education. Reestablishing links between universities and research and breaking down the political powers in scientific research has been more difficult than first thought.

Furthest along the track are Poland and Hungary, where universities and academies had managed to maintain their distance from the communist party. Furthest behind are Romania and Bulgaria. Romania is the only former communist country that has not tried to evaluate its research activities because of the extensive damage done by the Ceausescu regime. Bulgaria, in a similar but less severe situation, planned an evaluation last October but failed to reach a consensus on how to proceed.

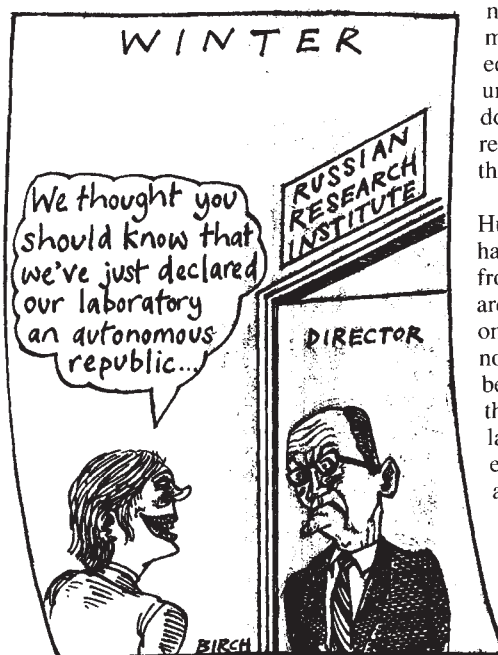
Between the extremes lies the former Czechoslovakia, whose decision to divide delayed the reform process on both sides. Czech research is relatively strong but academic restructuring has hardly begun. By contrast, Slovak research is weaker but its academic system has always been more liberal. The Czech Republic starts the new year with a new, government-directed research plan; Slovakia hopes to institute a science policy by April.

Science in the former Soviet Union probably faces a prolonged economic crisis. Many fear the disintegration of an infrastructure that once provided pockets of world-class research.

Central and Eastern European countries continue to call for foreign aid as short-term measures to help stem the flow of scientists to the West. The European Commission has recently allocated ECU55 million to fund cooperative projects during 1993. Individual institutes — and some individuals — in the East have offered help of various sorts, and the solidarity within certain close-knit international communities, such as astrophysics, has also meant practical support for some projects.

But these initiatives are dwarfed by an economic depression on a scale not seen for decades in Western Europe. In such circumstances, science and research will remain low on any government's list of priorities in 1993.

Alison Abbott



1990 was greeted with unsurpassed optimism; by contrast, the mood in 1993 could hardly be more different.

Holding up the best is the former East Germany itself. Universities and research centres, restructured along Western lines with formidable speed and determination, were relaunched on 1 January this year on the same legal basis as those in western Germany. In the past three years, the German Academy of Sciences and its institutes have been dismantled, but two new national research centres and two Max Planck Institutes (plus eight departments) have been founded. Nineteen applied research Fraunhofer Institutes set up in 1991 must meet their goals by the end of this year or face closing.

Although the stage is now set for a bright long-term future, it has not been an easy three years. Academic pay is still only 80 per cent of that in the west, and no-one knows the fate of the tens of thousands of scientists dismissed from overstuffed institutes during the ruthlessly enforced renewal process.

Although the methods have caused pain and resentment, the worst is now over. More intractable problems face other countries

Fast forward for gene therapy

Last month's approval by the Recombinant DNA Advisory Committee (RAC) of the National Institutes of Health (NIH) of three independent proposals to begin limited trials for the severe hereditary lung disease, cystic fibrosis (CF), heralds a new era in gene-therapy research for common hereditary diseases.

Until now, gene therapy has focused largely on expensive *ex vivo* approaches for rare diseases, including certain forms of cancer and immunodeficiency, although the RAC last year also approved ambitious proposals for brain and lung cancer. By contrast, the CF proposals call for the direct administration *in vivo* of the normal CF gene encapsulated within an adenovirus vector. Despite earlier concerns about the safety of such viruses and the efficacy with which they can be prevented from replicating, the RAC passed the three proposals unanimously. Many experts believe that such direct administration is necessary to realize the true potential of gene therapy in a clinical setting.

In the United States, at least 15 clinical gene-therapy trials are in progress, so far with mixed results. Efforts to treat rare genetic disorders such as adenosine deaminase (ADA) deficiency in two children and familial hypercholesterolaemia (FH) in an adult woman are proving effective. Indeed, James Wilson of the University of Michigan, whose team is treating the female FH patient with the low-density lipoprotein receptor gene, was recently granted permission by RAC to extend the study to four new patients. By contrast, studies using transformed tumour-infiltrating lymphocytes have come under fire (see *Nature* 360, 399; 1992) and many other trials have not yet yielded conclusive results. Research questions remain about the purification of viruses in large quantities and the targeting of specific populations of cells.

The next twelve months may see European researchers gaining ground. Two prominent European investigators, Dinko Valerio in the Netherlands and Claudio Bordignon in Italy, are already attempting to target stem cells to prolong the effectiveness of the treatment for ADA deficiency. In Germany, a proposal by Roland Mertelsmann to transfect the cytokine interleukin 2 into fibroblasts before using them as a vaccine for cancer patients was recently approved by the state of Baden-Württemberg (see *Nature* 360, 702; 1992).

In Britain, however, a proposal by Valerio's collaborators to begin trials for ADA deficiency, as well as a separate strategy for CF therapy from investigators in London, have suffered because the government has not yet set up a formal expert review board. Similar concerns among other European investigators regarding the

paucity of approved gene-therapy trials has prompted some to approach the RAC about the possibility of using it to examine European-based proposals.

With anxieties about the safety issues and effectiveness of gene therapy decreasing, 1993 could see a spate of new proposals as investigators turn to new diseases and alternative strategies to deliver healthy genes. Many other proposals are likely to be of the 'gene marketing' variety, in which researchers tag populations of cells such as bone marrow cells with a recombinant marker gene before transplanting them back into the host.

Kevin Davies

Recession grips industrial R&D

This year seems likely to be another gloomy 12 months for industrial research and development (R&D) as the global recession continues to depress spending and force cutbacks in research. One of the few bright spots is the pharmaceutical industry, which is weathering the recession with continued heavy investment in drug development.

Last year, even Japan, which maintained buoyant growth into the 1990s, found it could not escape the tide of recession. Major players in Japan's industrial research, such as the electronics and automobile industries, have been forced to freeze R&D budgets as their profits fell to all-time lows (see *Nature* 356, 93; 1992). The end of 1992 saw International Business Machines (IBM) announce a cut of \$1 billion in its planned outlay on R&D worldwide in 1993, the latest in a string of bad economic news that included last summer's announcement by British Petroleum that it will shed nearly half of its research staff in Britain and the United States (see *Nature* 358, 528; 1992).

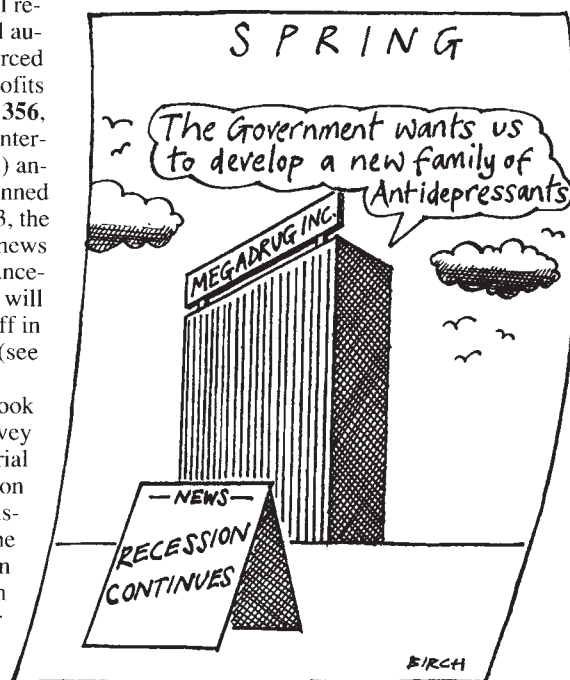
Prospects for the coming year look even grimmer. An end-of-year survey by the Washington-based Industrial Research Institute, an organization representing 260 prominent industrial companies, predicts that the recession will hit R&D spending in the United States even harder in 1993. More than a quarter (28 per cent) of the 141 companies replying to the survey say that they plan to cut their R&D budgets this year and even more (36 per cent) will spend less on equipment and R&D facilities. In Japan, the recession has prompted some bizarre economy measures, including switching off the heating during lunch breaks and eliminating business cards for researchers (see *Nature* 360, 98; 1992).

Despite a supplementary budget of more than ¥10,000 billion (US\$80 billion) introduced by the Japanese government a few months ago to boost the economy, the recession in Japan is expected to continue well into this year, and many observers predict that companies will be forced to slash their R&D budgets during the 1993 fiscal year, which begins on 1 April.

The recession is also affecting Japanese government research spending. Nevertheless, industry, including some non-Japanese companies, is turning to Japan's Ministry of International Trade and Industry (MITI) for support of long-term precompetitive research.

The situation is particularly severe in Europe because of the costs of reunifying Germany. In western Germany, 1,500 research positions will be shed in national research institutes and Germany has cut its funding to large-scale European projects, such as the European Space Agency and CERN (the European Laboratory for Particle Physics) (see *Nature* 360, 701; 1992). The development of a strong research base in Spain through increased government spending has also been delayed by the recession (see *Nature* 360, 502; 1992). And the leading government organization for applied research in the Netherlands, the TNO, has been told to find more money from industry because of cutbacks in government support (see *Nature* 360, 402; 1992).

The only bright spot in all the gloom is



the drug industry. As one of the directors of Japan's Takeda Pharmaceutical company puts it, "people have to continue buying drugs even in a recession". Glaxo, the British-based pharmaceutical group and one of the largest companies in the world, plans to boost its R&D spending from £600 million

(US\$900 million) to £720 million in 1993 to try to maintain its success in developing such popular drugs as its anti-ulcer medicine Zantac. And while most of Japanese industry has held down R&D spending, Japan's drug industry increased its spending on R&D in 1992 by 6 per cent, a rate well above inflation. Similarly, in the United States and Switzerland drug companies are maintaining heavy investment in drug development, exceeding 10 per cent of sales.

Another glimmer of hope is offered by the election of Bill Clinton as US president. Many observers expect Clinton to strengthen government support of industrial research, much as the Ministry of International Trade and Industry does in Japan. But any such policies are unlikely to have much effect on spending during 1993.

David Swinbanks

Agricultural biotech moves into spotlight

The year is likely to be the most critical yet for the agricultural biotechnology industry. This summer, all eyes will be on Calgene, Inc. of Davis, California, as it introduces its rot-resistant Flavr Savr tomato into the \$3.5-billion annual US market. A successful launch will also pave the way for the next wave of genetically engineered foods and make it easier for agricultural biotechnology companies to raise capital.

Roger Salquist, Calgene's chairman and chief executive officer, has become the industry's standard bearer in building consumer confidence in genetically engineered foods and doing battle with critics. A frequent foe is Jeremy Rifkin of the Washington-based Foundation on Economic Trends, who agrees that this year will also be critical for the foundation. He has promised to step up his 'Pure Food Campaign', and to lead a boycott of each new genetically engineered food until regulators implement pre-market testing and labelling of the foods.

All Calgene has actually done with its Flavr Savr tomato is to isolate the gene that codes for the polygalacturonase enzyme and reintroduce it into selected tomato varieties in the reverse or 'antisense' orientation. That change blocks the action of the enzyme, which causes softening of the fruit, by as much as 99 per cent, producing a better-tasting, longer-lasting tomato that can be harvested from the vine when ripe instead of ripening artificially. If Calgene's tomato is rejected by consumers, predicts Louis Da Gama, executive director of the UK BioIndustry Association (BIA), the idea of introducing novel genes into crop plants "will be dead in the water".

The Flavr Savr tomato is expected to be followed within a year by crops able to tolerate herbicide treatments, and within two to three years by crops resistant to insect attack. Eventually, companies hope to

develop families of genetically modified oils and fatty acids for use as foodstuffs and in industrial applications as biodegradable lubricants and hydraulic fluids.

At the same time, there is little reason for optimism among those working on novel animal technologies. Genetically engineered bovine somatotropin (BST), when injected into lactating dairy cows, can boost milk yields by 10 per cent or more but offers no benefit to consumers. Despite a clean bill of

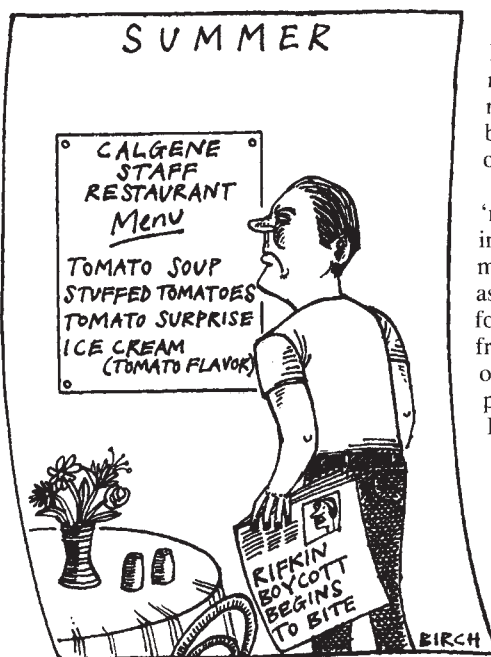
rise to discoveries, trained people and information that will be particularly useful to European industry". It is expected to cover research into plant development, with emphasis on hormones and related growth substances, in addition to environmentally friendly agricultural research, research to improve the quality of harvested products and the use of beneficial microorganisms.

Despite positive news last year for the agricultural biotechnology industry in the United States, a patchwork of regulations still exists in Europe. The Council of Ministers' directive on the deliberate release into the environment of genetically modified organisms, due to be adopted by member states in October 1991, has yet to be implemented by six of the 12 countries of the European Communities.

There is some progress expected on the 'novel foods' regulation, under discussion in the Council and in the European Parliament. The regulation sets out a means of assessing the safety of all novel foods and food ingredients, including those produced from or containing genetically modified organisms. The BIA opposes the current proposal, believing that it would place European industry at a disadvantage by going beyond international standards on safety, efficacy and quality.

Legislation is also under discussion in the EC for the labelling of foods produced by genetic modification. Last May, the US Food and Drug Administration gave preliminary approval to eliminating pre-market review and labelling of foods derived from biotechnology except in circumstances where levels of a naturally occurring toxicant have been increased, when an allergen not usually found in a plant has been introduced and where levels of important nutrients have been changed (see *Nature* 357, 352; 1992).

Diane Gershon



health from several review panels, the drug has become a prime target for consumer lobbyists who have raised questions about its safety and who say that it will put small farmers out of business.

Some expect a smoother regulatory ride in the United States for genetically engineered porcine somatotropin, which when given to pigs produces leaner meat. Philip Paxman, chairman of the European Trade Association for Advanced Animal Breeding, points out that a more efficient, more profitable animal is not enough. "There also has to be a social benefit", he says, such as the use of animals to produce human therapeutics or as model systems to study human disease.

Early this year, the European Commission (EC) is expected to start a significant research initiative in agricultural biotechnology involving more than 50 research laboratories. Managed jointly by the John Innes Centre in Norwich, England, and the Max Planck Institute for Breeding Research in Cologne, Germany, and funded initially for three years, the programme is intended to strengthen the European science base and make it a stronger international competitor. Richard Flavell, director of the John Innes Centre, says that the programme will try "to create a platform of research right across Europe that is seamless and that will give

Magnetic fusion proceeds with no end in sight

The inauguration of the Tokamak Fusion Test Reactor (TFTR) at Princeton University on Christmas Eve 1982 was described by the world media as the dawn of the era of fusion energy. Almost exactly a decade later — and 35 years after Sir John Cockcroft said (mistakenly) that he was "99 per cent certain" fusion had been demonstrated at Britain's Harwell laboratory — clean, commercially available energy remains a long-term dream.

This year, for reasons both scientific and political, fusion is again likely to make headlines. On the experimental front, two events stand out: the Joint European Torus (JET) at Culham, England, is expected to reopen in October after being closed for

upgrading, the same month that the TFTR plans to start its own experiments with tritium — the basic fuel for a commercial fusion reactor.

The main purpose of the JET modifications is to study ways of reducing the impact of impurities caused by interaction between the burning plasma and the wall of the tokamak. Such impurities seriously limit the length of ignition; their effects are expected to be reduced by the insertion of a pumped divertor, which traps the impurities by modifying the shape of the magnetic field at the plasma edge and allows them to be removed by a cryopump.

When the divertor studies will be completed by the end of 1994, JET will enter its final stage of full-power operation with deuterium-tritium plasma. Before that happens, however, the spotlight will have been turned back to Princeton's TFTR, where scientists hope to study for the first time the heating effects of alpha particles on the plasma contained by the magnetic field inside the tokamak. Equally important in political terms is the fact that the tritium experiments will allow the TFTR to reach power levels of 5–10 MW, surpassing the 2 megawatts achieved in 1991 by JET using 11-per-cent tritium injected into the deuterium plasma.

The experiments at JET and Princeton will provide results for what is most likely to be the next major goal for the fusion community, namely the construction of the \$6-billion International Thermonuclear Experimental Reactor (ITER). Progress on ITER has been slow but steady. The formal agreement setting the framework for cooperation was signed last summer in Washington by the four international partners (the United States, the European Communities, the Russian Federation and Japan). The next twelve months should see a continuation of the design engineering work and the building up of research teams at the three design centres at San Diego in California, Garching in Germany and Naka in Japan (see *Nature* **360**, 615; 1992).

Halfway measures

Banking on possible delays to ITER, a number of national initiatives have been put forward as intermediate steps. In the United States, the political focus will be on a proposal from Princeton to build a new Tokamak Plasma Experiment (TPX), a \$500-million project which, if approved, would begin operation around 2003 (see *Nature* **356**, 96; 1992). It would succeed the TFTR, which will be closed in 1994 because of the radioactivity caused by the use of tritium in the forthcoming series of experiments. After winning the backing of the US Department of Energy under the Bush administration, its fate now lies in the hands of President-elect Bill Clinton and his vice-president, Al Gore.

Across the Pacific, more discreet lobbying efforts are under way to persuade the Japanese government to follow up the

successful operation of its own fusion machine, JT-60, with a new steady-state tokamak. It is also seen as a stepping-stone to ITER. In Europe, as the main political focus shifts from JET to ITER, pressure is also likely to mount for funding for an alternative advanced machine, namely a stellarator, which physicists would like to build at the Institute for Plasma Physics in Garching.

Newspaper headlines in Europe, however, are likely to concentrate on a more pressing question, namely how to handle the consequences for scientists of the eventual closing of JET. The issue is of particular concern to British staff who, unlike colleagues from other European states, work under contract to the British AEA Technology, rather than to Euratom.

Networks enhance use of computers

Scientists can be expected to continue finding new ways to take advantage of improvements in the scope and speed of high-performance computers. As computing power becomes more crucial to national and regional economic security, two areas attracting the most attention are massively parallel processing, in which thousands of individual processors perform millions of tasks simultaneously rather than sequentially, and the linking of high-performance computers to networks.

The European Communities (EC) are being urged to begin a decade-long high-performance computing research programme using public and private funds amounting to ECU1 billion (US\$1.25 billion) a year. Part of the sales pitch, outlined in an October report by a European Commission panel led by CERN's Carlo Rubbia, is that there exists a "unique window of opportunity" for Europe to invent new hardware and software applications for the latest breed of massively parallel high-speed computers. The current research and development programmes, ESPRIT and RACE, have pushed the technology ahead and produced startup companies, but the report laments that larger companies from outside Europe often are quicker to exploit the markets created by new technologies (as Fujitsu Ltd hopes to do in acquiring the British computer maker ICL).

The name that Japan has given to its high-performance computing initiative — the Real World Computing project — reflects an emphasis on applications-orientated research, as well as a step back from the lofty goals of its less-than-successful predecessor, the now cancelled Fifth Generation project on artificial intelligence. Meanwhile, Fujitsu will attempt to move ahead of the pack with what promises to be the fastest supercomputer yet.

In the United States, which leads the world in massively parallel processing and

A report commissioned by the Commission at the prompting of the European Parliament has recommended that British scientists be given the same contractual status as their European colleagues (see *Nature* **357**, 270; 1992). But this will cost money, and implies a long-term commitment that the British government has so far not been willing to make.

The JET council meets in March to discuss its next move. Without a shift in Britain's position, the council is unlikely to be in a position to recommend a solution to the dispute. Union representatives are already warning that such an outcome could lead to a renewal of last summer's strikes. This looming dispute suggests that fusion may generate some heat this year, even if not the type originally intended. **David Dickson**

semiconductor research, the new Clinton administration is preparing to launch the second phase of a High-Performance Computing and Communications project to provide new supercomputer applications for health care, education and manufacturing. A key component on the project, launched in 1991, is the five-year effort to create a National Research and Educational Network using optical fibre cables to connect universities, major research centres, industry and possibly even homes.

Towards that end, the US Department of Energy and American Telephone and Telegraph Co. are improving the speed of such a network. One test scheduled for this summer will involve researchers using two Cray supercomputers to link 30 computers capable of transmitting 1 billion bits of information. That would be equal to sending 100 novels every second. The speed of the US Internet databank system will also be improved in 1993, from 1.5 million bits per second to a consistent 45 million bits per second, with laboratory tests in the mid-100 million to one-gigabit range.

Greater efficiency

Research into the field of high-performance computing will continue to focus on making supercomputers faster, but an equal challenge will be to make them work more efficiently. A central goal is rapid transmission of massive amounts of data in the form of moving video. The supercomputers seem up to the task, but technical problems inhibit networks from carrying such vast amounts of data. Optical fibre network switches so far are unable to detect and correct transmission errors at such high speeds. Regulatory logjams also plague the network in the United States, as telephone companies lack government permission to charge rate tariffs sufficient to cover the costs of transmitting large volumes of high-speed data.

In the realm of parallel processing,

computer engineers must also grapple with the problem of standardizing computer software in a way that manufacturers and businesses find appealing. One leading option is "high-performance Fortran", which adapts one of the oldest and most widely used computer languages to massively parallel supercomputers.

Another avenue is to combine the best features of parallel and traditional vector supercomputing, which use a much more efficient linear method of calculation. By September, Fujitsu plans to start selling a supercomputer capable of performing at 355 billion operating instructions per second, or gigaflops, that will use a combination of the two technologies and do in a minute what would take 10 days on a typical desktop workstation. Fujitsu claims the machine will be able to use existing software, while also having the performance advantages of parallel processing. Cray also plans to roll out a hybrid system that would allow users to

attach parallel processing features to their current supercomputers. And in June, Oak Ridge National Laboratory in Tennessee will begin using an Intel Paragon supercomputer that will operate at a peak of 150 gigaflops.

Further down the line, the Japanese project combines research on massively parallel machines with optical computing, using beams of light rather than electronic impulses to transmit data. The Japanese have urged researchers from around the world to join the project. Although cost-sharing might sound appealing at a time when EC scientists struggle with the uncertainties of a divided Europe and US scientists are weaning themselves from a shrinking defence department research budget, so far few companies have taken Japan up on the offer — another indication of the intense competition in the field.

Michael Mills

Hubble repairs to top active year in space

The US National Aeronautics and Space Administration (NASA) hopes that 1993 will bring an end to its biggest public relations headache of recent years as space programmes around the world prepare for 12 months of down-to-earth politicking as well as extraterrestrial exploration.

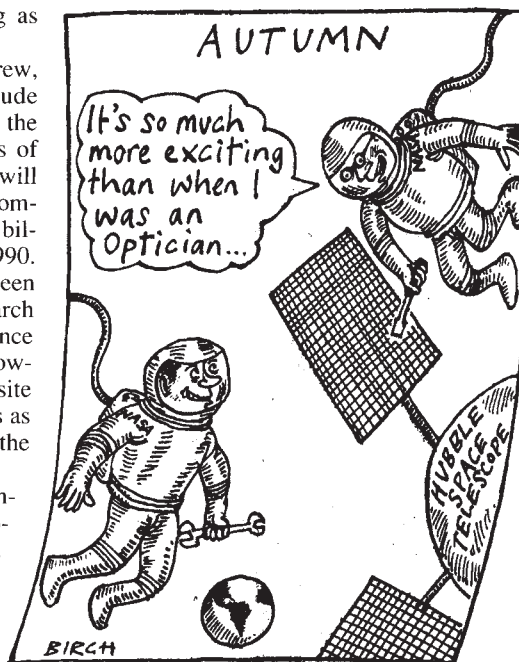
In December a seven-member crew, including European astronaut Claude Nicollier, will attempt to rehabilitate the Hubble Space Telescope. In a series of complex space walks, the astronauts will install a set of corrective devices to compensate for flawed optics in the \$1.5 billion orbiting telescope, launched in 1990. Even with the defect, scientists have been able to carry out a range of research projects, including searching for evidence of black holes in distant galaxies. However, the repair mission is a prerequisite for other such eagerly awaited efforts as measuring the rate of expansion of the Universe.

In April the same shuttle, the Endeavour, will try to retrieve the European Space Agency's Eureca satellite. Placed in orbit this past July, Eureca carries experiments to determine the effects of radiation and weightlessness on such objects suspended outside the satellite as photoelectric cells and plant spores. It is designed to be launched repeatedly after being plucked from orbit and returned to Earth.

In August, NASA's Mars Observer spacecraft is expected to begin orbiting the red planet. Meanwhile, the space agency's Magellan probe of Venus is expected to cease operations in May.

As usual, most new scientific and science-related spacecraft in 1993 will be launched on unmanned expendable rockets.

Geologists, oceanographers and others who use remote-sensing spacecraft to monitor events on the Earth's surface will benefit from two new satellites. In September,



France will launch its SPOT-3 remote-sensing spacecraft, replacing the ageing SPOT-2 satellite, and next month the US government had planned to launch its Landsat 6 remote-sensing spacecraft to replace its Landsat 4 and Landsat 5 spacecraft, which have exceeded their scheduled lifetimes. But a defect in the spacecraft's solar panels is expected to cause an indefinite delay.

Other significant scientific payloads include the launch in August of NASA's Sea

Wide-Field Sensor spacecraft to monitor chlorophyll levels and improve understanding of the global carbon cycle; the launch this month of Brazil's first homemade satellite, to relay environmental data, and next month's launch of Astro-D, Japan's X-ray satellite. The US Defense Department also plans to orbit a scientific spacecraft, known as Alexis, carrying both X-ray telescopes and an experiment to monitor interference with radio transmissions.

Several spacecraft will explore solar physics. NASA's Wind spacecraft, due to be launched next December, will work in concert with Japan's Geotail spacecraft launched in 1991. The Russian space agency's principal scientific launch this year will be a pair of spacecraft to monitor interactions between the solar wind and the Earth's magnetosphere.

Russia also has plans for three manned flights to dock with the Mir space station, including a flight in July featuring a three-week visit by a French cosmonaut. In November, a Russian cosmonaut is scheduled to fly on a US space shuttle mission, setting the stage for a visit in 1994 by a US astronaut aboard Mir.

However, despite all the activity on launch pads and control rooms around the globe, the most far-reaching developments affecting space may occur in legislative chambers and government offices. In the Commonwealth of Independent States, officials are trying to continue a vibrant space effort in the midst of continuing economic and political turmoil. Compounding the problem is squabbling among the three republics with the greatest interest in space: Russia, which builds most spacecraft and rockets; Kazakhstan, which is the launch and landing site for most missions; and Ukraine, which is home to some manufacturing and telemetry facilities. Simultaneously, the member countries of the European Space Agency are attempting to forge closer relations with the massive post-Soviet space apparatus in an era of tight budgets on both sides.

In the United States, the Clinton-Gore administration and a substantially new Congress are expected to rekindle longstanding political battles relating to space. For example, as a US senator, Vice President-elect Al Gore badgered reluctant Pentagon officials into approving access for environmental researchers to previously secret measurements made by US spy satellites. As chairman of the National Space Council, Gore may be able to make the military respond more quickly.

And in the US Congress, the planned space station Freedom is likely to be the most contentious issue in space policy. During the presidential campaign, Clinton and Gore voiced support for Freedom. But it is far from clear that Congress, faced with stiff competition for a tight budget, will hold to past agreements to build the \$37-billion orbiting platform.

Vincent Kiernan

Eighteen ninety-three and all that

J. L. Heilbron and W. F. Bynum

This year's commemorative cornucopia includes sunspots, the cotton gin and some stimulating acronyms. And a new award, 'the year of the anniversary', is introduced.

In earlier announcements in these pages, we have introduced 'the anniversary of the year', a concept as useful as it is obvious. This year we declare a more sophisticated concept, 'the year of the anniversary'. Like most conceptual breakthroughs in science, it rests on observation and compromise. The observation: in one and the same year, Nicolas Copernicus and Andreas Vesalius published books that were to transform astronomy and anatomy, respectively. The compromise: take the year of these revolutionary publications, 1543, as the unit of celebration, the 'year of the anniversary'. We may add in support of our choice that 1543 also saw the publication of Petrus Ramus's attack on Aristotle, *Animadversiones aristotelicae*; the arrival of the Portuguese in Japan; and the establishment of the first modern bureaucratic pile of papers, the Spanish state archives in Samancas. To some it may be a further recommendation that 1543 is a prime number.

As before, we exhibit our candidates for anniversarial attention in anti-chronological order, beginning at the first centennial and ending at the fifth; after which we make a few abbreviated suggestions taken from the declining twentieth century. Again we use 1.0¢ to indicate a centennial, 1.5¢ a sesquicentennial, and so on.

1893 (1.0 centenary)

Thermodynamicists will want to mark this year. In 1893 Josef Stefan, who deduced from experiments on glowing

wires that the radiant energy emitted from a hot body is proportional to the fourth power of its absolute temperature, and John Tyndall, who made the experiments, both died, independently. Their work was carried forward by Wilhelm Wien who, also in 1893, showed that the most intense frequency in the radiation is proportional to the temperature. Wien recognized that his 'displacement law', as the relationship between frequency and temperature was called, referred strictly only to a black body, to which, in 1884, Ludwig Boltzmann had restricted Stefan's law. Further to the defeat of heat, James Dewar (died 1923, 0.7¢) invented the Dewar flask, a vacuum-jacketed bottle for preserving cold bodies at low temperatures.

Anyone wanting to celebrate that *ex nihilo nihil fit* may wrongly seize on the experience of Henri Moisson, who announced 100 years ago that he had succeeded in making diamonds from charcoal. It appears, however, that his electrical oven could not achieve the temperature and pressure required to make gems from soot. Perhaps more electricity, which was no longer too cheap to meter (in 1893 Frederick Weston invented a gauge for measuring consumption of electrical energy), would have helped. Still, Moisson obtained something more valuable than diamonds from his oven: it is mentioned together with his experiments on fluorine, which he was the first to isolate and, in collaboration with Dewar, the first to

solidify (in 1903, 0.9¢), in the motivation for his Nobel prize in chemistry awarded in 1906.

Three growth industries began in 1893. The first four-function calculator, the aptly-named 'millionaire', first went on the market; the first issue of the *Physical Review* appeared; and Robert Diesel described his engine. The growth of the computer and automobile industries, and of scientific publishing, may be inadequately indicated by the enlargement of the *Physical Review* by a factor of 400, from 120,000 words in 1893 to around 48,000,000 last year.

At the Columbian Exposition held in Chicago in 1893, which missed the commemoration of Columbus's landing in the New World by only 1 part in 400, the Krupp industries exhibited a gun whose 1,000-kilogram shell could pierce a metre of armour at a distance of one kilometre. The world no doubt rejoiced to know that the gun was intended not for sale, but for advertisement; which could be done the more easily and effectively since that very year, 1893, C. V. Boys invented a way to photograph artillery shells in flight. When Krupp's kilogram shell landed it made an impression locally indistinguishable from any of the three classes of earthquakes then introduced into science by Rudolf Hoernes, who thereby reduced by 25 per cent the four categories proposed by John Milne in 1883 (1.1¢). Hoernes's three: accompaniments of volcanic eruptions, collapses of void spaces within the Earth, and slips along fault lines. Residents of

Mary Evans

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Histology of the cerebral cortex, an optical telescope, the barometer and the foundations of anatomy represented by (left to right) Camillo Golgi, Claude Chappe, Evangelista Torricelli and Andreas Vesalius.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Hypnotic technique — Jean-Martin Charcot demonstrating at la Salpêtrière.

northern California, and seismologists everywhere, were given something to think about when Andrew Lawson, a Scottish geologist who settled in Berkeley, described the San Andreas Fault.

In 1893, Jean-Martin Charcot, after two decades' work on hypnotism, died. His critics thought he must have been half asleep to have taken up the subject in the first place, but some of Charcot's therapeutic methods were already being transformed by Sigmund Freud and Josef Breuer. Their paper of that year, on the psychic mechanisms in hysteria, was the first salvo of a brief partnership that launched Freud on his road to the unconscious and left Breuer wondering why his collaborator thought so many psychiatric deviations were rooted in sex.

Meanwhile, Karl Pearson took up other kinds of deviation. His lectures in late 1892 and early 1893 on the 'laws of chance' used the familiar (card drawing, coin tossing) to illustrate the recondite. Scientists more than punters latched on to his methods of calculating standard deviation, introduced in the lecture of 31 January 1893 to replace older formulations based on fuzzier concepts.

1843 (1.5 centenary)

In February 1843 a comet could be seen around lunchtime, close to the Sun. This compelling appearance was but one of several reasons to celebrate a sesquicentennial of the Sun this year. In 1843, John William Draper, a chemist turned photographer, took the first solar portrait in the infrared. Heinrich Schwabe, a druggist turned astronomer, announced the ten-year sunspot cycle — news that went against received opinion, which held that the spots came and went capriciously. Some fifty years later (1.0¢), Edward Walter Maunder worked out from recorded spots that the Sun's complexion was relatively clear between 1645 and 1715 (the 'Maunder mini-

mum'), and so confirmed the concept of a 'little ice age'.

Communication might also qualify for commemoration. This year 150 years ago Siegfried Becker of Vienna prepared the world's first bottle chart, on which he inscribed the itineraries of 119 bottles tossed here and there into the ocean; the longest journey, of 17,000 kilometres, took 33 months. He did much better than last year's anniversarian of the year, Christopher Columbus, whose bottle containing a copy of his report of his discovery of the Indies, thrown overboard off the Azores in 1493 (5.0¢), has not, to our knowledge, been recovered. Recently, current plotters obtained a possible substitute for bottles when a storm attacked a ship carrying 40,000 Nike gym shoes across the Pacific. All went overboard. The belief that not a sole had survived was shattered by an armada of Nikés arriving on the coast of Oregon later in the year. Look for more in Hawaii and Japan this year.

Further to communication, Alexander Bain (died 1903, 0.9¢) prematurely invented a telegraphic fax machine: William F. Cooke introduced the first signal blocks, on the Eastern County Railroad; and Isambard Kingdom Brunel built the first tunnel under the Thames. Charles Wheatstone's famous bridge, though of the correct date, does not belong among the year's contributions to communication. It is a useful technique for measuring relative electrical resistance. In both its usefulness and relativeness it contrasted with the absolute method of electromagnetic measurement then being introduced into physics by Wilhelm Eduard Weber.

If more material for sesquicentennials is needed, in 1843 Marc Antoine Augustin Gaudin invented the alloy of platinum-iridium that, in a slightly different proportion from his, later served as the material of the standard meter; Carl Gustav Mosander isolated the rare-

Mary Evans

earths erbium and terbium; William Rowan Hamilton invented non-cumulative algebra; and Gustave-Gaspar Coriolis, whose fictitious force makes hurricanes, and Charles Mackintosh, whose raincoats do not protect against them, died.

Births could have been safer in 1843 had doctors heeded Oliver Wendell Holmes's little essay on the contagiousness of puerperal fever. Not many of them did. Among those who survived the rigours of birth were Sidney Irvine Smith, who dredged crustaceans off the New England coast, and David Gill, who mapped the stars from Cape Town; Camillo Golgi, who studied the histology of the cerebral cortex, and David Ferrier, who was more interested in its functions; and Robert Koch, who needs no introduction.

1793 (2.0 centenary)

Further to communication (see under 1843), in 1793 Claude Chappe interested the National Convention of the brand-new French Republic in his optical telescope. His first installation, of 1790–91, had been torn down by a Parisian mob indifferent to science on the suspicion that Chappe was signalling to the royalist enemies of the revolution. With the backing of the Convention, he set up 20 stations between Paris and Lille. Skill and patience were required: the sending of a single sign over the entire distance took six minutes. It would have taken hours to announce and explain what may be the most famous decision of the Convention regarding science: its order, in August 1793, to close the world's leading scientific organization, the Paris Academy of Sciences, as a feudal relic unnecessary in a revolutionary society.

The Paris Academy had compromised itself by, among other things, its arrogant dismissal of inventions submitted to it for its approval. Would it have recognized in the year of its demise the merits of the first important invention made in the new United States, Eli Whitney's cotton gin? The United States did, and rescued Whitney from bankruptcy in a way still practised by government — with a contract for firearms.

While the American arms-maker Whitney was ginning cotton, the Protestant pastor Christian Sprengel was admiring geraniums. Believing that the hairs on the petals of *Geranium sylvaticum* have some purpose, he devoted years to scrutinizing the pollination of flowers by insects. His *The Discovered Secrets of Nature in the Construction and Fertilization of the Flower* came out in 1793, complete with a thousand engravings from his own drawings and the conclusion that the dichogamy he observed so frequently showed that God did not intend that plants commit incest.

Charles Darwin was later to learn German to read books like Sprengel's and to reach different conclusions.

1743 (2.5 centenary)

No European city of the eighteenth century was fully respectable without a learned academy. In 1743 Philadelphia fleetingly enjoyed this respectability in the form of the American Philosophical Society, promoted by Benjamin Franklin (anniversarian of the year for 1990). The British colonies had not yet attained the degree of civilization needed for science and the society failed for want of men, money and interest. It obtained a permanent lease on life 25 years (0.25¢) after its putative founding, when what remained of it joined with a struggling society of letters in 1768. Another omen of the improvement of science in America dating from 1743 was the birth of Thomas Jefferson.

While philosophy was aborting in America, it was flourishing in France. Two hundred and fifty years ago this year, Jean le Rond d'Alembert, a founding raised on mathematics, published his *Traité de dynamique*, which contains the first foundation of what physicists call d'Alembert's principle for reducing dynamics to statics. D'Alembert died in 1783 (2.1¢), having outlived his great rival Alexis Claude Clairaut (born 1713, 2.8¢), whose father, a mathematics teacher, did his part for multiplication by siring 20 children. Clairaut helped newtonianize the Paris Academy of Sciences and also the Marquise du Châtelet, the translator of Newton's *Principia*; and in 1743 he brought out his own important development of Newton's ideas, the *Théorie de la figure de la terre*. In it, Clairaut demonstrated the shape of the Earth from hydrostatic principles, developing along the way the mathematics of what would later be called the potential; and also an approach to capillarity that later inspired Laplace to create his celebrated theory of the rise of liquids in narrow tubes. Laplace survived, even thrived, during the revolution, but two savants born in 1743 did not: Antoine-Laurent Lavoisier and the Marquis de Condorcet, whose violent deaths in 1794 allow an unusual accelerated anniversary observation. Their births can be celebrated this year, their deaths commemorated next year. The Royal Society has already noted that its longest-serving president, Joseph Banks, was also of the class of '43.

1693 (3.0 centenary)

To the ranks of newtonian fellow travellers deserving recognition this year should be added John Wallis (died 1703, 2.9¢), the second volume of whose *Opera mathematica*, published in 1693, contains the first complete statement in

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

A view of Moordaener's Bay, New Zealand: engraving by I. Gilsemans from Tasman's journal.

print of Newton's method of fluxions. Wallis acted from fear that foreigners otherwise would claim the glory of the mathematical discoveries of the 'greatest man who ever lived', Isaac Newton, who, at this very moment, in 1693, was sunk in the blackest depression. Perhaps Newton's reading and working in alchemy, which had occupied much of his time since completing the *Principia*, deepened his despondency; for try as he might, he could not make his alchemical knowledge make sense. By the end of the year he had recovered himself and abandoned the art; believers in the power of rationality might therefore schedule a major celebration this year.

While Newton was breaking down, Edmund Halley was cutting up. The man who had prompted Newton to write out the *Principia* had hit on a simple way to compute the relative areas of land and water on Earth. Taking the gores from a good terrestrial globe, he cut out the continents and larger islands and weighed the scraps. Halley was also curious about his own mortality, constructing in 1693 the first life table, with "its Moral, its Physical and Political Uses". Based on data from Breslau, it showed him that, were he prepared to cross the English Channel as a 36-year-old, he could expect to live another 24 years. He stayed in London and made it another half-century (died 1743, 2.5¢ on the old calendar).

1643 (3.5 centenary)

An important contribution to the mapping on which Halley relied had been made 50 years earlier by Abel Jansen Tasman, the average date of whose voyage across 5,000 miles of uncharted South Pacific we take as 1643. He set sail from Batavia in August 1642 on commission from the Dutch East India Corpora-

tion. He discovered the Friendly Islands in January 1643, having first found New Zealand (where the natives were not friendly) and Van Dieman's land (now Tasmania), and before stumbling over the Fiji and Tonga islands. He circumnavigated Australia without observing it, thus disproving the notion, derived by European knowledge of its west coast, that it joined the huge suppositious southern continent of terra incognita, and showing that it is possible to discover much without seeing anything. Tasman's employers judged his trip worthless since, as their evaluation put it, "he had found no treasures or matter of great profit". He had no worthy successor until the voyages of Captain Cook in 1773 (2.2¢).

The year 1643 was prolific in information about what is not there. About the time that Tasman showed explorers the disconnection between Australia and Antarctica, Evangelista Torricelli, a student of Galileo's student Benedetto Castelli (died 1643, 3.5¢), may have had the idea of the barometer and the theory of the void space above the mercury. But just as the force of the vacuum and the Horror vacui fell as a result of Torricelli's invention and explanations, so has the date 1643 for their creation succumbed to the labours of historians. Anyone wishing to celebrate the traditional date is welcome to do so, and would have the additional opportunity of commemorating Gasparo Berti (died 1643), whose demonstration of a water barometer inspired Torricelli's work; but the informed anniversalist will wait until next year.

1593 (4.0 centenary)

Shakespeareans may wish to rejoice at the end of the Bard's apprenticeship and the beginning of his creative maturity;

and anglers may be cognizant of the birth of Izaak Walton; but we predict that there will not be many quadricentennial celebrations of science this year.

1543 (4.5 centenary)

Nicolas Copernicus was on his deathbed when he received the first copy of his exposition of the heliocentric system, the revolutionary *De revolutionibus corporum coelestium*. That was just as well. His editor, Andreas Osiander, had added a preface to the text without bothering to identify its author and without telling Copernicus. This bit of pseudepigraphy or reverse plagiarism made Copernicus appear an instrumentalist and even a positivist. Osiander's preface exhorts the reader to consider heliocentric astronomy as a *jeu d'esprit*, a hypothesis, a mathematical artifice, and warns that unless he, the reader, understands the truth that in astronomy there is no truth, he "will leave this subject a greater fool than when he entered it".

Speaking of fools, Peter Ramus made out, also in 1543, that Aristotle's logic was silly. He also said an unkind word about Aristotle's physics. The attacks earned him some unphilosophical responses from fellow philosophers and an edict from the King of France prohibiting him from teaching or writing philosophy. In fact, Ramus attacked not the long-dead Aristotle, for whom he professed the deepest respect, but the relatively more alive modern aristotelians, his opponents in the schools. After four years of silence, Ramus resumed the attack when the ban on his philosophizing was lifted in 1547.

By that date, Andreas Vesalius had forsaken the theatres of academe for the antechambers of the court of Charles V. The gluttonous Emperor gave him plenty of experience in treating bowel disorders, but nothing Vesalius did after 1543 came close to the verbal and visual brilliance of his *De humani corporis*

fabrica, which gave anatomy its compass. He was 29 when it was published, reminding us that anatomists as well as mathematicians may peak early. According to Halley's life tables, Vesalius's death in 1564 on the way back to Spain from the Holy Land was a decade premature: a warning either against taking pilgrimages or, perhaps, against leaving academic life for more lucrative employment.

1493 (5.0 centenary)

It would be churlish to leave last year's anniversarian of the year high on the Ocean Sea. Those who want to prolong the columbian quinquicentennial can observe his return to Spain; the publication of his discovery; the sailing of his second expedition, bent on colonization and evangelization; and the division of the spoils of exploration between Spain and Portugal by the Pope.

1943 (0.5 centenary)

A transposition of figures returns us to living memory and the age of abbreviations and acronyms, without which modern science is inconceivable. According to the *OED* (we should rejoice that it was not called the *Dictionary of Oxford English, DOE*), the word acronym entered our language in 1943, together with one of the most popular acronymic inventions.

Jacques-Yves Cousteau, then in the French underground movement, wanted to spend more time under water. Accordingly, with the engineer Émile Gagnan, he developed the aqualung, which allowed him freedom of the water without diving suits and long tubes. Among Cousteau's predecessors in inventing self-contained

Mary Evans

underwater breathing apparatus were Leonardo da Vinci and Stephen Hales. Cousteau's scuba gear was more practical than theirs. It has proved to be equally convenient for marine biologists, underwater archaeologists, treasure hunters and sun seekers. After the war, Cousteau founded a marine research group in the French navy before becoming one of the most effective of scientific and ecological communicators.

Abbreviations are

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Christopher Columbus's triumphal entry into Barcelona on returning from his first voyage.

more common than acronyms. Two entered currency and consciousness in 1943. Adrenocorticotrophic hormone was already known, but ACTH was finally obtained pure from sheep pituitaries by Choh Haoh Li (born 1903, 0.8¢), working in Herbert Evans's laboratory in Berkeley. Three years earlier, he had isolated luteinizing hormone (LH), adding another abbreviation to endocrinology, a science especially rich in them.

On 16 April 1943, the Swiss chemist Albert Hofmann had to leave his laboratory early, after accidentally absorbing some of the chemical he was working with. He published a description of the extraordinary mental symptoms and hallucinations he experienced and expressed the hope that the drug, lysergic acid diethylamide, might have some use in the investigation and, possibly, treatment of mental disorders. It certainly generated an enormous literature — scientific, social and legal — though its early promise never materialized. As a cultural phenomenon, LSD might better appear under 1968, were we inclined to notice anything so recent in this year's anniversarial harvest. □

W. F. Bynum is at The Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BE, UK; J. L. Heilbron is vice-chancellor of the University of California at Berkeley, Berkeley, 200 California Hall, California 94720, USA.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Peter Ramus delivering his thesis against Aristotle, which led to a ban on his philosophizing.

Adventures with a vulnerable knee

Knees are intricate heavy-duty joints that are now, like many other parts of the body, partially replaceable. A knee surgeons' handbook now published shows what lies ahead.

THAT people who suffer surgery develop an awesome respect for their surgeons is well known. The psychoanalysts have a word for it: the transference. What follows is an acting out (another analysts' concept) of such a transference, which takes what may be considered the unusual form of a book review. The book, a substantial tome by any standards, with more than 800 pages, is called *Knee Surgery* (Dunitz, London; £145.00). The editors of a symposium volume are Paul M. Aichroth from the Westminster Hospital, London, and W. Dilworth Cannon Jr, from the University of California, San Francisco. The circumstances of the transference are these.

Just under 20 years ago, the first of the editors removed a cartilage (otherwise a meniscus) from my right knee in an old-fashioned operation, from which it took several months to recover. When, in the summer of 1991, I twisted the same knee and sensed the same old pain, I telephoned asking for advice, arrived for an appointment at what had become the "knee unit" at a private hospital and was offered a disconcerting choice: "Would you like me to do it tomorrow or a week tomorrow?"

It emerged that this informality was a testament to the efficacy of arthroscopy — the technique for operating within the knee through three small holes, one of which carries the arthroscope itself, a light source and associated imaging optics which is routinely hitched up to a video recorder. (The equipment and the operation itself were developed in Japan in the early 1960s.) Editor Cannon says that the operation is now the most common of all in North America, and that "few techniques are left for open surgery". In the event, I remember spending a day sleeping off the anaesthetic and watching the abortive coup against Gorbachev before walking out to catch a taxi to the office.

So it was natural to expect as pleasurable an outcome when I banged the same knee against a piece of furniture last October. The X-rays showed a crack near the inner rim of the kneecap, but this time there was no arthroscopy, only pain-killers, physiotherapy, instructions on how to walk (not hop) on crutches, CNN's version of the US election campaign and then the same nonchalant taxi ride to the office. But then I twisted the knee. Nobody knows why the muscles of that leg went into tremor, but I was carted off by ambulance for the second time and told that I had better turn the hospital room into an office. (We never

hooked up the modem, though.)

This is how the transference happens. One is told in plain language what is going to happen to one, in this case for at least the week ahead, and one's astonishment at one's willingness to accept the prescription without protest, proof though it may be of one's pathetic eagerness to be cured, is best rationalized by supposing that the instructions come from a supremely confident being. At one stage during that long week, Aichroth casually let slip that "We're doing total knee replacements now".

Indeed they are. Michael A. R. Freeman from the London Hospital gives an account of 21 years of experience with an artificial knee originally developed in collaboration with the Department of Mechanical Engineering at Imperial College, also in London. The trick is to replace the load-bearing surfaces in the knee joint with inert materials without interfering with too many of the attachments within the knee of the tendons and ligaments that hold it together, and which ensure that it will flex without becoming dislocated.

To those who are not surgeons, the procedures are really quite breathtaking. People describe the way of using mechanical jigs that guide the saw-blades that provide flat surfaces onto which to seat the artificial materials. Freeman prefers high-density polyethylene for the lower bearing surface, which is seated (and usually cemented) in the upper part of the trimmed-down tibia, and a cobalt-chromium alloy for the upper surface, wrapped around the head of the femur. But there is plenty of room for argument, and for improvement of design. Some favour titanium rather than the cobalt-chromium alloy. Others are working with knee joints in which the stability is provided by mechanical constraints, as in a hinge. Some even use the word "hinge" . . .

Nobody offered me a knee replacement, and if they had I hope I would have summoned up the strength to say, "Not yet". In the end, when the knee had failed fully to recover, there was another arthroscopy, debris was excavated from behind the kneecap (bones bleed when broken, one forgets), another miracle cure and another taxi to the office. Ten days in hospital, that took.

I wish now that I had listened more carefully to the advice I was given before I left. Take it easy! Go to physiotherapy regularly, and preferably every day! Do not give up the crutches until your knee feels well! Keep icing that knee! The trouble lies in the layman's sense of what a cure means. Am I

well or am I not? If the latter, I will stay in bed, if the former, I will walk. So I bought a walking-stick in a farm supply shop and went on my travels.

To be fair, Aichroth and Cannon is full of the need for retraining the knee after an operation of any kind, under the rubric of rehabilitation. Their associate Dipak V. Patel, from the Wellington Hospital, London, has laced their volume with careful statistical accounts of people's recovery from the several procedures they describe.

Outsiders by now will not be surprised to learn that the evaluation of knee function is in itself a serious issue: indeed, according to a contribution by Aichroth and Patel, there have been so many worthy but uncoordinated attempts at the evaluation of knee function that a committee called the International Knee Documentation Committee has been set up to rationalize them. Like the surgeons they are, they endorse the view that evaluation should not tempt people "to substitute bad accounting for good judgement".

In reality, the fun in this volume is in the sense of daring it repeatedly conveys. There are, for example, enthusiastic accounts of the usefulness of nuclear magnetic resonance in visualizing the soft-tissue structures in the knee. Then there is the business of dealing with the serious injuries of the stabilizing ligaments to which sportspeople are prone. Mathematicians will be intrigued by the insufficiently explicit accounts of where best to re-fix ligaments to the neighbouring bones, but surgeons (for whom the book is written) will be more taken with tales of how to replace them altogether. Plainly the knee is well on the way to being replaceable.

So why my regret at having taken so quickly to a walking-stick? After an overseas journey, I remembered that I was due for a check-up. Slightly querulously, I complained that the knee still hurt, and was told to have patience. That evening, I slipped on the kitchen floor and broke my right femur just above the knee. From where I had fallen, I told the surgeon what had happened. He said he'd look in at the hospital in a couple of hours (by which time it was 2.30 a.m.). No doubt sympathetically, he mentioned that the death rate from broken femurs during the First World War was 90 per cent. The following day he installed a stainless steel plate, complete with an array of screws. Drooling over an X-ray of my thigh, a physiotherapist said "who would have thought anybody could have put that together again?" The knee-cap, as it happens, also seems to be mending. **John Maddox**

Tunguska comes down to Earth

H. J. Melosh

FEW natural events this century have excited as much popular and scientific interest as the gigantic explosion that rocked the Siberian taiga near the Tunguska river on the morning of 30 June 1908. Second only to the Loch Ness monster and Bigfoot in the popular science press, the Tunguska explosion has moved otherwise reputable scientists to propose explanations ranging from antimatter meteorites to mini black holes to near-critical fissionable masses. One theory popular in 1930s Russia attributed it to the explosion of a nuclear-powered (!) spacecraft. And the UFO explanation seems prevalent in some quarters even today. However, serious scientific study has been converging on a more prosaic explanation wrapped up by Chyba and colleagues on page 40 of this issue¹.

At its simplest, this explanation holds that the Tunguska explosion was caused by the fall of a large meteorite that broke up and deposited its energy in atmospheric blast waves before reaching the ground. The essentials of this picture were established by L. A. Kulik, who made the first on-site investigations of the explosion in the years 1927–39. Kulik attributed some boggy depressions near the explosion site to meteorite impacts, a proposal that was later discredited. No impact crater or large meteorite fragments were ever found at Tunguska.

Kulik's work was refined by E. L. Krinov², who proposed that the explosion was created by a comet. The idea of a cometary impactor was strongly supported by more recent work³, although the density of the required comet is very small, 0.01–0.001 g cm⁻³. Compared to the density of roughly 0.6–1.0 g cm⁻³ reported for comet Halley^{4,5}, this would make the Tunguska object decidedly unusual. However, Chris Chyba, Paul Thomas and Kevin Zahnle now argue that a full consideration of the dynamics of a meteorite traversing the atmosphere shows that the Tunguska explosion is fully compatible with the entry of a roughly 30-m diameter meteorite of the common stony class.

Previous workers considered a low-density comet necessary because of the peculiar pattern of energy deposition: the arrangement of the flattened trees, microbarograph and seismic records all suggest a nearly point-like release of 10–20 megatons equivalent of TNT in the atmosphere about 10 km above the surface. Although the 'butterfly' pattern of the tree-fall does suggest some energy released along a sloping line, it is clear that most of the energy was deposited in

a small region near the terminus of the giant meteor's flight. The only way for a meteorite (or comet) both to deposit the observed energy and to stop in the atmosphere at the observed height is if it possesses a very low average density. This is a very general result from the conservation of energy and momentum, and seems very difficult to get around.

Chyba and colleagues' innovation is to

craters). In Chyba and colleagues' picture, the low average densities are achieved by fragmentation and dispersion of an object that was originally of normal density.

The idea that meteorites are fragmented by aerodynamic forces is not new. Apart from the bursts and flare-ups plainly visible in many bright fireballs (see figure), strewn fields of meteorites such as the 1947 Sikhote-Aline fall or the Henbury craters in Australia attest to atmospheric breakup of incoming meteorites. Quinn Passey and I examined this process in some detail 10

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Near miss: on 10 August 1972 this brilliant fireball made a 1,500-km track as a small asteroid dipped into the Earth's atmosphere over the western United States, passing just 58 km above Montana, at 15 km s⁻¹. Estimates of the object's diameter vary widely, from 4 to 80 m. If the larger estimates are right, and if it had approached at a steeper angle, it might have caused an atmospheric explosion comparable to Tunguska's. The photo above was taken at Grand Teton National Park in Wyoming.

include the effect of aerodynamic forces on the body. They show that these can fracture it and spread out the fragments, greatly increasing its overall atmospheric drag and so increasing the rate at which energy was deposited into the atmosphere. This process proceeds nearly catastrophically, as increasing drag increases the deceleration, which in turn spreads the fragments faster. In addition, the atmospheric density rises exponentially with decreasing altitude. These effects mean that bodies of the appropriate size will ultimately deposit most of their energy in a small region near the end of the trajectory (objects much smaller than 30 m in diameter either burn up or reach terminal velocity in the upper atmosphere, whereas the largest objects penetrate to the surface and form hypervelocity impact

years ago^{6,7}, as did Grigorian and others in the former Soviet Union^{8–11}. However, it was not until Zahnle studied the origin of the 'dark shadows' on Venus^{12,13} that a fragmentation model was coupled to the equations governing meteor flight in the atmosphere. Chyba, Thomas and Zahnle apply this model to the Earth, computing the energy deposition from fragmenting meteorites and comparing the results to the Tunguska explosion. They examine the entry of meteorites starting with 15 megatons of kinetic energy and with strengths associated with the known meteorite classes — the strong irons, ordinary stony meteorites and fragile carbonaceous chondrites. Several estimates for comet strengths are also used. The iron meteorites hardly fragment at all and so strike the ground at high speed, as other

irons in this size range have been observed to (Meteor Crater, Arizona, was made by an iron object with approximately this energy). The comets and carbonaceous chondrites deposit their energy too high. Only the ordinary stony meteorites create an explosion resembling that at Tunguska.

Substantial progress has thus been made in reducing the Tunguska explosion from the realm of the near-miraculous to a natural, although rare, occurrence. Instead of an implausibly low-density comet, the Tunguska projectile was more likely a representative of the most common class of meteorites. This picture leaves a few questions unanswered, however. One of the more notable phenomena associated with the Tunguska explosion were the widely reported 'light nights' observed over Eurasia for the first few nights after the explosion. Proponents of the cometary impact theory could point to the possibility of light reflected from parts of the comet tail that missed the Earth, but stony meteorites do not possess tails. Instead, Chyba and colleagues revive an earlier suggestion that the light nights were due to sunlight reflected from high noctilucent water-ice clouds. They estimate that the rise of the Tunguska 'fireball' of air heated by the explosion injected a large quantity of water into the upper atmosphere, which could have been transported over Europe by high-altitude winds.

There is still work to be done on the physics of Tunguska-like explosions. Chyba and colleagues' model of meteoroid fragmentation and dispersal is relatively crude. However, the physics of meteoroid fragmentation in the Earth's atmosphere is similar to that needed to study impact cratering on Venus, so there is reason to hope that further work on this problem will be forthcoming from both planetary and meteoritical scientists. □

H. J. Melosh is in the Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721, USA.

Sticking to the point

Paul W. Kincade

INFLAMMATION involves the cooperation of many different types of cells and the interaction of myriad different proteins. A possible role for one of these — the cytokine macrophage inflammatory factor (MIP-1 β) — is provided by Tanaka *et al.* in a paper on page 79 of this issue¹. They find that MIP-1 β triggers one stage in the sequence of events that leads to lymphocytes adhering to the walls of capillaries before migrating out of the blood stream to sites of inflammation in tissues.

A consensus model emerged last year in which at least two different classes of adhesion molecules are required for leukocyte extravasation². The first step involves members of the selectin family which slow the flow of leukocytes through postcapillary venules, promoting their transient adhesiveness to the endothelium. The second involves integrins, which convert the transient interaction to a strong adhesion. The 'adhesion triggers', which were thought to influence the activity of integrins, remained incompletely defined. MIP-1 β is one of a number of cytokines involved in the initial stage of leukocyte migration, and Tanaka *et al.* now show that MIP-1 β could be such a trigger. It acts on the CD8⁺ subset of T lymphocytes — cytotoxic T cells — which respond by increasing the adhesiveness of their integrin VLA-4 for one of its ligands — VCAM-1. The problem of delivering this stimulus efficiently to cells that are still within the blood stream is apparently solved by tethering the cytokine to endothelial cell-surface proteoglycans — immunostaining of tissue sections has revealed the cytokine in that location. Thus, lymphocytes that slow sufficiently to interact with the immobilized cytokine will be encouraged to activate integrins, which in turn mediate firm adhesion by recognition of ligands on the vessel wall.

Extravasation of cells from the blood stream is essential for normal lymphocyte migration and inflammation, and similar mechanisms operate in the metastasis of some tumours. The importance in this process of members of both the integrin and selectin families of cell adhesion molecules (CAMs) has been known for some time. The two families of CAMs work in sequence³. Transient interactions of rapidly moving leukocytes with endothelium are mediated by one of three animal lectins (L-selectin, E-selectin or P-selectin) and their carbohydrate ligands. This favours multiple collisions between the two cell types and the typical 'rolling' behaviour of neutrophils, which precedes their flattening against

the vessel wall. This latter step is clearly mediated by integrins.

The functions of some CAMs are regulated by changes in their levels of expression on the cell surface. Integrins, however, are constitutively expressed and their functions are regulated in other ways. 'Activation' states of cells expressing integrins and the extracellular divalent-cation concentration dramatically affect the specificity and avidity of integrins for their ligands. This active metabolic process, termed inside/out signalling, involves phosphorylation/dephosphorylation steps as well as interaction of cytoplasmic integrin tails with the cytoskeleton or other molecules⁴.

The integrin VLA-4 (α_4, β_1) studied by Tanaka *et al.* is found on lymphocytes, monocytes, eosinophils and blood-cell precursors, and their experiments indicate how VLA-4 might be selectively activated on a particular T lymphocyte subpopulation.

There is growing awareness of the functional interdependence of proteoglycans, cytokines and CAMs^{5,6}. Certain growth factors and cytokines, including members of the so-called 'intercrine' family to which MIP-1 β has been assigned, have consensus sequences for binding to glycosaminoglycans, such as heparan sulphate on cell-surface proteoglycans. Cytokines thus immobilized can be sequestered, protected from degradation and presented to neighbouring cells. Heparan sulphate is also a critical component of the receptor for at least one cytokine fibroblast growth factor — and can be recognized as a ligand by at least one CAM (N-CAM). Tanaka *et al.* emphasize that some intercrines are localized on endothelial cells in postcapillary venules. These 'pro-inflammatory' cytokines are made in abundance by stimulated macrophages and other cells⁷ and it is now clear that MIP-1 β can be 'pro-adhesive' for at least one T-cell subset. A mechanism is provided for tethering it to the luminal surface of blood vessels, thus preventing its being washed away by the blood stream.

Tanaka *et al.* used heparan sulphated cell-surface antigen CD44 to coat plastic dishes and to immobilize MIP-1 β for presentation to T lymphocytes. It is not clear that CD44, rather than some other transmembrane proteoglycan, bears the cytokine on endothelial cells. It was, however, an interesting choice of molecule for a number of reasons. CD44 is clearly involved in a variety of processes, and is widely distributed on endothelial cells and blood cells. It has recently been found to have many protein isoforms,

1. Chyba, C. F., Thomas, H. J. & Zahnle, K. J. *Nature* **361**, 40–44 (1993).
2. Krinov, E. L. *Giant Meteorites* (Pergamon, Oxford, 1966).
3. Turco, R. P. *et al.* *Icarus* **50**, 1–52 (1982).
4. Sagdeev, R. Z., Elyasberg, P. E. & Moroz, V. I. *Nature* **331**, 240–242 (1988).
5. Peale, S. J. *Icarus* **82**, 36–49 (1989).
6. Passey, Q. & Melosh, H. J. *Icarus* **42**, 211–233 (1980).
7. Melosh, H. J. in *Multi-ring Basins* (eds Schultz, P. H. & Merrill, R. B.) 29–35 (Pergamon, New York, 1981).
8. Grigorian, S. S. *Kosmicheskie Issledovaniya* **17**, 875–893 (1979).
9. Levin, B. Y. & Bronshten, V. A. *Meteoritics* **21**, 199–215 (1986).
10. Ivanov, B. A. *et al.* *J. geophys. Res.* **97**, 16167–16181 (1992).
11. Ivanov, B. A., Basilevsky, A. T., Kryuchkov, V. P. & Chernaya, I. M. *J. geophys. Res.* **91**, 412–430 (1986).
12. Zahnle, K. J. *J. geophys. Res.* **97**, 10243–10255 (1992).
13. Melosh, H. J. *Nature* **358**, 622–623 (1992).

including one that is involved in tumour metastasis⁸. As an additional source of variation, some cells glycanate CD44 with heparan sulphate and/or chondroitin sulphate. Again, this may correspond to the aggressive behaviour of tumour cells. CD44 can itself mediate cell adhesion and one mechanism involves its ability to recognize hyaluronate. As in the case of integrins, cells expressing CD44 must be in an appropriate 'activation' state for the molecule to bind hyaluronate⁹. The findings reported here call attention to another potentially important role for CD44, that of a docking site for regulatory cytokines.

Biological redundancy is a lesson frequently being taught by 'knock out' mice, and we could soon learn that some of the many CAMs, proteoglycans and

cytokines are not essential. In the meantime, it is interesting that representatives of several large families of molecules must cooperate in a complex and seemingly vital biological process. □

Paul W. Kincade is in the Immunology Program, Oklahoma Medical Research Foundation, 825 NE 13th Street, Oklahoma City, Oklahoma 73104, USA.

1. Tanaka, Y. *et al.* *Nature* **361**, 79–82 (1993).
2. Butcher, E. C. *Cell* **67**, 1033–1036 (1991).
3. Lawrence, M. B. & Springer, T. A. *Cell* **65**, 859–873 (1991).
4. Hynes, R. O. *Cell* **69**, 11–25 (1992).
5. Nathan, C. & Sporn, M. J. *Cell Biol.* **113**, 981 (1991).
6. Flaumenhaft, R. & Rifkin, D. B. *Curr. Opin. Cell Biol.* **3**, 817–823 (1991).
7. Oppenheim, J. J., Zachariae, C. O. C., Mukaida, N. & Matsushima, K. A. *Rev. Immun.* **9**, 617–648 (1991).
8. Gunther, U. *et al.* *Cell* **65**, 13–24 (1991).
9. Lesley, J. *et al.* *J. exp. Med.* **175**, 257–266 (1992).

GONDWANALAND

Models come in from the cold

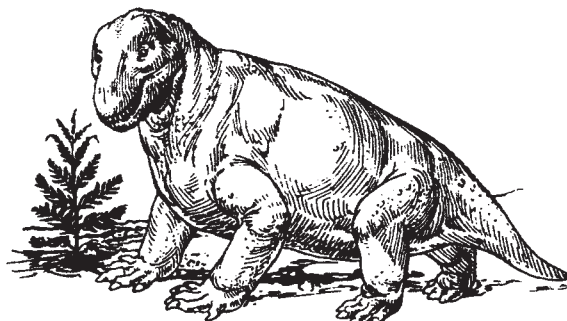
Alfred M. Ziegler

FAST lakes up to 2,000 km long might have had a drastic moderating effect on the climate of the ancient supercontinent Gondwanaland, Yemane argues on page 51 of this issue¹. She takes climate modellers to task for ignoring geological data that indicate a temperate climate for Gondwanaland in the late Permian (250 million years ago) and for publishing model results indicating wild seasonal temperature fluctuations, more severe even than in Siberia today. And she is sending geologists back to their drawing boards to prepare better maps of the supercontinent. The lakes seem to be the element overlooked in both disciplines.

Numerical modelling developed for the prediction of current climate change has given a fresh impetus to the study of the climates of past periods when the world's geography was radically different. Geology provides the only data against which the models can be checked. But problems arise because not only are the models imperfect, but so are the geological data, and the geographical reconstructions are often too crude even for the coarse computer grids used in modelling.

Part of the trouble is that several factors conspire to produce changes in geography, including plate tectonics, sea-level variation and mountain building. Moreover, the geological record of these effects is very incomplete. The thoughtful palaeogeographer must relate the available information to some concept of the processes involved, but the end result may include more imagination than fact. Particularly difficult is tracing out shorelines of both lakes and epicontinental seaways, making uncertain any

estimate of the size of such bodies. Yemane starts from lake deposits, which have been known for many years, and which she suggests are erosional remnants of lakes, up to 2,000 km long and 500 km wide, that existed in the deep interior of Gondwanaland (which comprised the modern southern continents).



Sturdy evidence: the 5-foot high *Moschops*, an ungainly mammal-like reptile and herbivore, apparently enjoyed life in the late-Permian interior of Gondwanaland, to the discomfort of modellers who predicted extremes of climate. The suggestion that *Moschops* migrated to avoid the harshest conditions is scorned by Yemane, who doubts their ability to cover great distances. (From ref. 7.)

In my view, Yemane has taken a realistic, indeed conservative, approach in mapping these late-Permian lakes; it is possible that the Karoo Basin of South Africa was even larger, and connected to the sea by way of the Parana Basin of Argentina and Brazil. It has been shown that this was the case earlier in the Permian², and an extension of this sea some 2,500 km into Antarctica has recently been proposed³. Traditionally, all these deposits were classed as terrestrial or fresh water, but a few fossils are now known that allow one to infer the

presence of brackish water and this implies an interconnected inland seaway of considerable proportions. Parts of this basin were filled in with sediments, but some sectors could have retained a marine connection as late as Yemane's map interval. In sum, southern Africa and adjacent areas may have been more water than dry land, in marked contrast to the current palaeogeographic maps which show it all as land.

Lakes and seaways have an ameliorating effect on the climate. Technically, water has a lower specific heat capacity than rocks, but because it is a fluid it can be heated to depths of 50 m or more during the warmer months. This energy is then available for redistribution during the winter, when the land surface has long been frozen. Accordingly, the Gondwana lake deposits assume an importance disproportionate to their size and must be identified and searched carefully for clues to their original extent. In fact, this is one of the goals⁴ of the International Geological Correlation Program.

The most direct evidence of past climates is through the plants and animals that actually had to endure them, as well as through the sedimentary products of these environments. Evidence for a cool temperate climate in the Permian of southern Africa (40–55° S) is abundant and well summarized by Yemane. For comparison, London (51° N) and Chicago (42° N) experience such conditions today. One may add to her list the fact that a number of the lakes, as well as the inland brackish sea⁵, contain organic rich shales. Such deposits would not be expected where surface temperatures approach 4 °C (water's temperature of maximum density) in winter, because of turnover and oxygenation of the bottom due to density reversals⁶. In fact, the existence of these deposits points more to the warm temperate part of the spectrum.

If the direct line to palaeoclimate is through the geological record, then one might well ask, why perform the numerical studies in the first place? The computer runs are expensive but relatively fast, whereas the assembly of worldwide fossil datasets is labour intensive and the information is inevitably geographically limited. Palaeontologists as a group seem to be preoccupied with taxonomic or evolutionary studies, and are generally unwilling to apply themselves to other fields, except for the occasional symposium article. This is odd because geography and cli-

mate provide the context in which the biological patterns are determined, and one might well wonder how these factors can be ignored.

Despite these problems, the next few years will see a substantial improvement in our knowledge of past climates. Global palaeogeographic mapping is progressing slowly but surely here at the University of Chicago, with particular attention being paid to reconstructing the heights of mountain ranges and the widths of ancient seaways. Current reconstructions treat orography sketchily, and seriously underestimate the dimensions of water bodies. Improvements in the maps will show some continental dispositions that were even more different from the present than we realize, and the resulting climate reconstructions will also look less like the present.

Improvements in the atmospheric scientists' general circulation models will include grids with finer resolution and will couple them with ocean-circulation models. The interaction of the ocean and atmosphere is especially important because sea surface temperature anomalies help to drive the very atmospheric wind systems that create them in the first place. Oceanic circulation is difficult to understand, let alone to model, although crude systems are now available to climatologists. In most published reconstructions, the sea surface temperatures have been specified initially and the end results may be influenced by a rather arbitrary choice.

Testing of the numerical models will be mainly with palaeobotanical datasets currently being assembled. Even if the models can be validated, this testing process will be necessary on a continuing basis because of uncertainties in the palaeogeography. The Permian example of Yemane demonstrates how important it is to integrate the results of diverse approaches. Only when this is accomplished for successive periods of the Earth's history can the full spectrum of global change be revealed. □

Alfred M. Ziegler is in the Department of the Geophysical Sciences, The University of Chicago, Chicago, Illinois 60637, USA.

Violent motions in galactic core

James M. Moran

WHERE zoologists can tag animals with radio transmitters to track their movements, or atmospheric scientists can follow hurricanes with balloon-borne transmitters, astronomers have to rely on astrophysical chance to provide beacons to reveal large-scale motions. Using bright maser emission from traces of water vapour, Nakai, Inoue and Miyoshi have apparently found flows as fast as $1,000 \text{ km s}^{-1}$ within 2 parsecs (6.6 light years) of the nucleus of the nearby active galaxy NGC4258. The significance of the maser emission is in what it may tell us about the dynamics of the energy source driving the galaxy's large-scale jets. Nakai *et al.* suggest, on page 45 of this issue¹, that if the maser signals are not from energetic bipolar outflows from the galactic nucleus, or from gas orbiting a black hole, then they may be revealing the presence of dense clouds of ionized gas which scatter the maser emission. Each possibility is intriguing.

More than 1,000 masers have been found in our Galaxy since their accidental discovery 27 years ago². Originating from trace molecules such as hydroxyl or water vapour in gaseous astrophysical condensations, these narrow-line radio beacons are as bright as gas heated to temperatures of 10^{15} K or more. They are generally located in the envelopes of newly formed stars or evolved stars nearing the end of their life cycle. What these two types of stars have in common are copious winds, which shed material at rates of $10^{-8} - 10^{-3}$ solar masses a year for brief periods. The condensations form in these flows, and some of their molecular constituents are pumped by infrared radiation from the host star to stimulate maser radiation. Typical flows have speeds of around 30 km s^{-1} , but the more energetic ones reach 300 km s^{-1} .

These motions can be tracked by very-long-baseline interferometry (VLBI) and, with the addition of Doppler information, the three-dimensional velocity field can be measured³. The distance to the host star can also be estimated trigonometrically by comparing the angular velocities measured by VLBI with the line-of-sight velocities from the Doppler shifts of the lines in much the same way as a casual observer might determine the distance to an airplane by comparing its angular and linear velocities. In evolved stars, the masers in different molecular species, such as SiO , H_2O and OH form in nested shells where the species-dependent excitation conditions are favourable.

Over the past decade, masers a million

times more powerful have been discovered in the nuclei of about 50 active galaxies⁴. The galaxies are generally very bright at infrared wavelengths, possibly owing to bursts of star formation, but otherwise their 'megamasers' seem to have little in common with galactic masers. The most distant known megamaser is in IRAS20100-4156, which has a redshift of 0.129 corresponding to a distance of about 500 megaparsecs⁵. VLBI measurements⁶ of NGC3079 show that most of its masers are spread over a region 0.1 parsecs in diameter, about the same as in clusters of newly formed stars in our Galaxy.

The second most intense water-vapour maser known was found in NGC4258 during a search of nearby infrared galaxies by Claussen and Lo⁷. Its luminosity is equivalent to 100 Suns, a remarkable amount of radiation from a single transition of a trace gas. Nakai, Inoue and Miyoshi have now re-examined this galaxy. Their idea was to search for extreme velocity components in the galaxy that might be evident in the maser spectrum. Until recently, such a project would have been herculean, owing to the limited capabilities of spectrometers and narrow-band radio amplifiers. But Nakai, Inoue and Miyoshi had new tools: a wide-band, low-noise amplifier and a powerful spectrometer having 16,000 channels and the capacity to analyse a 300-MHz chunk of the spectrum (corresponding to a velocity range of $4,000 \text{ km s}^{-1}$) at one time.

With this equipment on the 45-m telescope at Nobeyama, they readily detected the known emission lines near the galaxy's systemic velocity of 450 km s^{-1} and, in addition, two clusters of lines offset from the systemic velocity by $\pm 1,000 \text{ km s}^{-1}$. What is striking about the spectrum is the wide gap between the masers at the systemic velocity and the outliers, and its remarkable symmetry. They also report preliminary VLBI observations that show all the masers to be coincident to within 2 parsecs.

What are these high-velocity lines telling us about NGC4258? One possibility is that they are not kinematic in origin at all. Several spectral artefacts can be ruled out, including the Zeeman effect (due to magnetic fields), the Stark effect (due to electric fields) and emission from unidentified molecules. The Zeeman effect would require an astonishing magnetic field of 10^5 gauss, as the H_2O molecule is nonparamagnetic. The stimulated Stark effect would need an unrealistic brightness temperature of 10^{23} K to achieve the requisite electric

1. Yemane, K. *Nature* **361**, 51-54 (1993).

2. Oelofsen, B. W. in *Gondwana Six: Stratigraphy, Sedimentology, and Paleontology* (ed. McKenzie, G. D.) 131-138 (American Geophysical Union, Washington, 1987).

3. Collinson, J. W. & Miller, M. F. in *Gondwana Seven Proceedings* (eds Ulbrich, H. & Rocha Campos, A. C.) 217-231 (Instituto de Geociencias, Universidade de Sao Paulo, 1991).

4. IGCP Project 324 *Global Palaeoenvironmental Archives in Lacustrine Systems* (contact Gierlowski-Kordesch, E., Univ. of Ohio).

5. Cole, D. I. & McLachlan, I. R. in *Gondwana Seven Proceedings* (eds Ulbrich, H. & Rocha Campos, A. C.) 379-390 (Instituto de Geociencias, Universidade de Sao Paulo, 1991).

6. Demaison, G. J. & Moore, G. T. *Am. Assoc. Petrol. Geol. Bull.* **64**, 1179-1209 (1980).

7. Colbert, E. H. *The Age of Reptiles* (Norton, New York, 1965).

field strength. There are spectral lines of HDO and HO^{37}Cl that would occur at about the observed frequencies, but those transitions have never been seen in astrophysical environments.

Stimulated Raman scattering has been proposed to explain large frequency shifts, using NH_3 as the intermediary, but the frequencies involved in this case would require multiple scattering of high order, which seems quite implausible⁸. Nakai, Inoue and Miyoshi suggest that Raman scattering of the photons from the water-vapour maser in a foreground plasma could explain the observed frequency shifts. An ionized cloud with a density of 10^8 cm^{-3} would have a plasma frequency of around 75 MHz, about the required amount. Such dense interstellar ionized regions are conceivable, but the geometric constraints required to obtain both a boost and a cut in frequency are difficult to realize.

Nakai, Inoue and Miyoshi also discuss two kinematic models. In the first, the masers trace the gas in a bipolar jet emanating from the nucleus; because the galaxy is inclined by 72° , the total velocity could be $3,000 \text{ km s}^{-1}$ if the jet lies along the galactic rotation axis. In the second, the maser emission comes from an accretion disk spinning around a black hole at a velocity of $1,000 \text{ km s}^{-1}$. A 1-parsec disk would require there to be a central mass of 2×10^8 solar masses to maintain keplerian orbital motion. The galaxy has a large-scale jet over 10 kiloparsecs long that appears prominently in the radio images of the galaxy's synchrotron emission. The angular and velocity structures have been measured in great detail by Cecil *et al.*⁹, who traced at least three braided strands in the jet. They postulate that the braids could originate either with a pair of orbiting black holes or with an accretion disk around one black hole with parameters close to those described by Nakai *et al.* If the masers are associated with the inner part of this jet, the water-vapour molecules would have to be accelerated gradually to avoid dissociation, in which case it is puzzling that no emission is seen in the gap between the systemic velocity and $\pm 1,000 \text{ km s}^{-1}$.

The definitive test of the maser models will undoubtedly come from VLBI observations, which should be able to

distinguish between rotation, expansion and no motion within a few years. Intercontinental VLBI can produce images with resolution of 200 microarcseconds at 1.3 cm wavelength; a transverse velocity of $\pm 1,000 \text{ km s}^{-1}$ will produce angular motions on the order of 30 microarcseconds a year, greater than the demonstrated sensitivity³ of a few microarcseconds a year. The fine details of the

motion will be discernible by VSOP and RADIOASTRON, two radio astronomical VLBI satellites expected to be launched by Japan and Russia, respectively, around 1995. □

James M. Moran is at the Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA.

EVOLUTIONARY BIOLOGY

The design of animal signals

Mark Pagel

ONE could be forgiven for thinking that the study of animal signalling consists principally of applications of Weber's law of just noticeable differences. Consider the problems confronting a male frog trying to attract a mate: he must design his 'song' to exploit the properties of acoustic signals, of the female frog's ear, and of the environment in which he will broadcast his song. But, as revealed at a recent conference*, he must also



A bluehead wrasse (*Thalassoma bifasciatum*) male showing its 'chase' colours. Is this a 'conventional' or 'honest' signal of intent (see text)? (Courtesy of Marian Dawkins.)

confront a world in which signal receivers have perceptual biases, hidden preferences and sensory systems burdened by history. If this isn't enough, the presence of other males trying to attract the same female means that it will usually pay for a male to make his signals showy, perhaps even frivolous or reckless. Evolutionary biologists, having taken important strides in both methodology and theory, now know nearly as much about signalling as the frogs, and in some cases more.

Animals produce a variety of signals which they use to threaten, to beg, to impress, to warn, to inveigle and so on. *Ex cathedra* scientists often smugly criticize the field biologists who study such signals for relying on anthropocentric techniques: how do we know that a bright yellow fish looks bright yellow to

those fish? The critics are going to have to get out of their chairs. An account by H. Römer, (Karl-Franzens University) of a 'biological microphone' made up of a live bushcricket with electrodes poked into its auditory nerves, and strapped atop a brass pole, was worthy of a cyborg-punk novel and made for good science. Connect this sentient bio-phone to a tape recorder, play another bushcricket's acoustic signal, and one can listen to what a bushcricket hears. Carry the bushcricket-cum-tape recorder around the forest and one can make a precise map of a signalling bushcricket's acoustic territory. Play several signallers and one can find out whether the listener can discriminate among them or even single one out — the so-called 'cocktail party phenomenon'. Bushcrickets can do both, but attend selectively to the loudest of

the competing inputs. This is important for testing ideas relating to mate choice and sexual selection; how females (in this instance) might choose among males on the basis of their 'songs', and how males should design their songs in response to the perceptual capabilities and preferences of females.

Female preferences are part of a theoretically sophisticated but contentious area of animal signalling — sexual selection. Darwin proposed that many of the secondary sexual characteristics of males, such as the long tails of birds of paradise, or the complex song repertoires of robins, evolved in response to female choice of mates. Theoretical models of sexual selection assume that the male trait and the female preference coevolve, leading to the exaggeration of the male trait and the evolution of female preference. Males benefit by being attractive to females and females benefit either from having sons with the trait (fisherian runaway), or by getting

1. Nakai, N., Inoue, M. & Miyoshi, M. *Nature* **361**, 45–47 (1993).
2. Elitzur, M. *Astrophysical Masers* (Kluwer, Dordrecht, 1992).
3. Gwinn, C. R., Moran, J. M. & Reid, M. J. *Astrophys. J.* **393**, 149–164 (1992).
4. Henkel, C., Bann, W. A. & Mauersberger, R. *Astr. Astrophys. Rev.* **3**, 47–90 (1991).
5. Staveland-Smith, L. *et al.* *Nature* **337**, 625–627 (1989).
6. Haschick, A. D. *et al.* *Astrophys. J.* **356**, 149–155 (1990).
7. Claussen, M. J., Heiligman, G. M. & Lo, K. Y. *Nature* **310**, 298–300 (1984).
8. Boyd, R. W. *Pub. astr. Soc. Pacific* **89**, 141–146 (1977).
9. Cecil, G., Wilson, A. S. & Tully, R. B. *Astrophys. J.* **390**, 365–377 (1992).

* The Evolution and Design of Animal Signalling Systems, London, 28–29 October 1992. Proceedings will appear in the *Philosophical Transactions of the Royal Society of London*.

good genes or some other resource from the father.

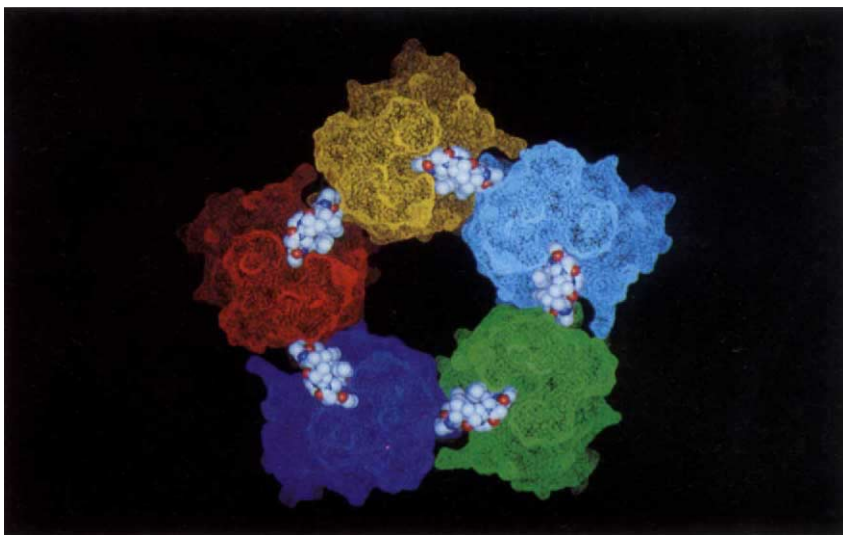
There is some evidence, however, that females may have 'pre-existing preferences' for particular traits, and that males may evolve merely to exploit females' sensory biases¹⁻³. Female frogs of the species *Physalaemus coloradorem* prefer the 'whine and chuck' calls of males from the sister species *P. pustulosus* to the 'whine' only calls of males of their own species (M. Ryan, University of Texas, Austin), even though they will not have heard them before. Phylogenetic studies suggest that the preference is ancestral to both species. It is significant that the hidden preference is for a qualitatively different call; a preference for a supernormal or exaggerated form of the call that they usually hear would be expected on theoretical grounds from the runaway, good genes, or better resources views of sexual selection.

Lucky will be the first *P. coloradorem* male to produce a call with chucks: theoretical models show that this trait could spread through the population even without any coevolutionary change in female preference⁴. Whether 'pre-existing bias' represents a new explanation for sexually selected traits or a special case of existing models is controversial. In *P. pustulosus* the larger males produce calls with chucks of lower frequency. Females not only prefer these calls, but have more of their eggs fertilized when they mate with the larger males¹.

Research that assigns genetic algorithms the task of learning discrimination tasks suggests that all perceptual systems have inherent and unpredictable hidden biases for stimuli that they have never seen (A. Arak, Archway Engineering, UK). To illustrate this point, consider an insect pollinator that has only two plant species in its environment, A and B, and prefers species A to species B. Flowers of species A have long petals, those of B are short, and the insect discriminates between the two flowers solely on this basis. Perfect discrimination would result if the insect's sensory system responded to A and B flowers with a step function: very low levels to B flowers, very high levels to A flowers, and very low levels to points in between. But an inverted 'U' shaped function that was low at B, higher at A, but highest in between would discriminate equally well between A and B, but would contain an inherent and never-tested sensory bias for flowers of intermediate size.

In principle, an infinite number of discrimination functions that yield hidden biases can be constructed. Such biases, if widespread in nature, would circumscribe the notion of an evolutionarily stable strategy (ESS)⁵ when applied to discrimination (Arak). All ESS

Decameric drug sandwich



THE IMMUNOSUPPRESSIVE drug cyclosporin A (CsA), a peptide fungal product that has been used for over a decade to prevent rejection of organ transplants, works by binding to members of the family of cyclophilin proteins. The drug-receptor complex mediates its immunosuppressive effect by interfering with the T-cell signal-transduction cascade, thereby preventing T-cell activation, a critical step in the immune response that causes transplant rejection. In investigating the X-ray structure of crystalline cyclophilin-cyclosporin A complexes, G. Pflügl *et al.* found the remarkable pentameric supermolecule shown above (cyclosporin, solid spheres; molecular surfaces of the five different coloured cyclophilin molecules shown as shrunk polygons using the program VIEW, M. Sanner, Sandoz). As the authors show on page 91 of this issue, these pentamers associate in pairs, sitting on top of one another, to form a ten-membered disk-like structure; the 'decameric sandwich'. The secondary structure elements of the cyclophilin A molecules, the β -sheets and α -helices, are on the upper and lower faces of the disk and the CsA-binding loops and ten CsA molecules are sandwiched between them (the picture illustrates the inner face of one half of the sandwich). The biological significance of the pentamer and the decameric sandwich is as yet unclear, but there are hints that they may be more than an artefact of the crystallization process. The X-ray structure, as well as revealing the intriguing supramolecular organization of the cyclophilin-cyclosporin complex, sheds further light on details of the interaction of the peptide drug and its receptor, independently analysed with NMR by Y. Thériault *et al.* on page 88. This information will be of great use in the design of CsA analogues.

G.R.

systems are limited, implicitly, to the range of stimuli considered, but this work makes clear the potential to jump out of such a system. At least temporarily, that is: for a jump out of an ESS system to be maintained, there must be an advantage to both the signaller and the receiver. Perhaps tellingly, the genetic algorithm's biases were for supernormal stimuli — it preferred bigger or longer instantiations of the stimuli which it had evolved to discriminate — rather than intermediate or qualitatively different forms. It is still early, but these results combined with the large number of selective forces acting on signalling systems (for example, most discrimination systems will be exposed to more than two forms) may greatly reduce the number and form of hidden biases.

Amotz Zahavi⁶ (University of Tel Aviv) carried a remarkable idea out of Israel and into the rest of evolutionary biology in the middle 1970s, an idea that after a short period in the wilderness has

now gained wide acceptance. Zahavi's idea was that when signals are used to persuade, as in threats or sexual attraction, signalling systems will evolve around costly, wasteful signals. Thus, the 'stot' of a Thompson's gazelle while being chased by a predator — as it runs it jumps up into the air holding its legs out straight — is an utterly wasteful, frivolous and even reckless thing to do. But this very wastefulness, ironically, assures the predator of the gazelle's ability to flee: stotting is an 'honest' signal because only the best quality gazelles can afford to 'handicap' themselves by stotting (see ref. 7). Similar logic has been applied to explain a diverse set of animal signals, and may help to explain why some people spend as much on a wristwatch as most of us would spend on a new car. Theoretical models of the handicap principle⁸⁻¹⁰ have assumed implicitly that receivers can perceive signals — such as the height of a stot — without error. New theoretic-

cal work¹¹, however, shows that the handicap principle works even when this assumption is relaxed (A. Grafen, University of Oxford). Receivers may also make use of several signals simultaneously in an attempt to 'average out' the error in their perceptions of each.

How pure are 'honest' signalling systems? Costly signalling systems could, in some circumstances, be vulnerable to cheats who form the signal but do not pay its costs (M. Dawkins, University of Oxford). Insects with warning coloration, such as ladybirds, and even some birds¹², sequester toxins and so may be signalling honestly to potential predators that they are unpalatable (but see ref. 13). However, 'automimic' ladybirds, which produce the warning colour but do not pay the price of sequestering the toxins, could exist at low levels. But cheats do owe their existence to honest systems.

The term 'conventional signals' has been applied to those signals for which the costs of the signal are not part of the message¹⁴. Much debate has focused on whether conventional signals can be used to persuade (see figure). Birds may use badges of status (D. Harper, University of Sussex), such as a spot on their breasts, to signal their position in the status hierarchy. A smaller, less costly, badge may be used to signal higher status. But even though the meaning of the signal (status) is not related to its cost, the actions required to enforce status will be, and so badges of status may in this sense prove to be 'honest'. Perhaps conventional signals commonly exist in cases where there is no conflict of interest between the two parties? Not necessarily: although a controversial point, genetically identical (save for mutations) somatic cells in the same body may ensure the reliability of signals passed between them by making those signals costly (Zahavi). □

Mark Pagel is at the School of Mathematical Sciences, Queen Mary and Westfield College, University of London, London E1 4NS.

Sensing calcium in rod cells

James B. Hurley

THE discovery of cyclic GMP-gated cation channels in the rod and cone photoreceptors¹ of vertebrate retinac unveiled an important and elegantly simple type of sensory transduction mechanism. Light-stimulated hydrolysis of intracellular cGMP is linked to inhibition of Na⁺ and Ca²⁺ influx by direct interaction of cGMP with plasma membrane cation channels. Recent reports²⁻⁴ demonstrate that Ca²⁺ regulates the affinities of these channels for cyclic nucleotides both in photoreceptors and in olfactory sensory neurons. Writing on page 76 of this issue⁴, Hsu and Molday report that the affinity of photoreceptor channels for cGMP is regulated by the ubiquitous Ca²⁺-binding protein calmodulin.

Electrophysiological studies of how light affects ion currents flowing across photoreceptor plasma membranes have defined important characteristics of the photoresponse. This has challenged biochemists to identify molecules and reactions that can explain those characteristics. Quite a few pieces of the puzzle of phototransduction have been identified, but few of them have been placed definitively into a picture that fully explains photoexcitation, recovery and adaptation. Hsu and Molday have identified an important new piece to be fitted into this puzzle.

Here is a popular version of how phototransduction works, based on what we have found out so far. Two well known second messengers, cGMP and Ca²⁺, have intricately intertwined roles in generating the photoresponse⁵. Light stimulates hydrolysis of cGMP through the actions of rhodopsin, transducin and a cGMP phosphodiesterase. In the presence of sufficient intracellular cGMP, as there would be in darkness, cGMP-gated cation channels allow Na⁺ and Ca²⁺ to enter the cell. Light-stimulated cGMP hydrolysis reduces this channel activity. Consequently, the intracellular concentration of Ca²⁺ falls as it is pumped out of the cell by a Na⁺/Ca²⁺-K⁺ exchanger⁶.

Ca²⁺ is an important regulator of certain vertebrate phototransduction enzymes. Lowering the concentration of free Ca²⁺ within the submicromolar range, as occurs in the cell following a light flash, stimulates cGMP synthesis by photoreceptor guanylate cyclase⁷. Low Ca²⁺ concentrations also reduce the light sensitivity and accelerate deactivation of photoreceptor cGMP phosphodiesterase activity⁸. Both Ca²⁺ effects were recently reported to be mediated by members of a newly identified family of Ca²⁺-binding proteins. The roles of these

proteins in signal transduction are only beginning to be understood. Initial reports^{9,10} suggested that one of these proteins, bovine recoverin, is a Ca²⁺-sensitive activator of guanylate cyclase and that another, frog S-modulin, regulates the light sensitivity of cGMP phosphodiesterase. But a report to appear in a forthcoming issue of *Neuron* raises important questions about the relationship between recoverin and guanylate cyclase. Gray-Keller *et al.*¹¹ have found biochemical and electrophysiological evidence that recoverin and related proteins do not promote recovery, but instead prolong the photoresponse.

Hsu and Molday⁴ show that Ca²⁺ in photoreceptors has an additional mode of action that is mediated by calmodulin. By measuring the cation permeability of vesicles derived from photoreceptor membranes, they found that Ca²⁺-calmodulin interacts with the rod photoreceptor cGMP-gated channel to raise its *K*_{1/2} for cGMP from 19 μ M to 33 μ M cGMP. As the cGMP-dependence of the channel is cooperative, this produced as much as a sixfold effect of calmodulin on channel activity at low cGMP concentrations. The calmodulin effect occurred between 50 and 30 nM Ca²⁺, which is probably within the range of free Ca²⁺ concentrations in a photoreceptor⁶. Only calmodulin produced this effect; it was not mimicked by several other Ca²⁺-binding proteins, including recoverin.

These findings prompted Hsu and Molday to identify photoreceptor proteins that interact with calmodulin. Not surprisingly, they found that the primary calmodulin target in photoreceptor membranes is the cGMP-gated channel complex. Calmodulin affinity chromatography purified two polypeptides of relative molecular mass 63K and 240K which had been previously shown to constitute the channel complex. Blots of purified channel preparations probed with radioactive calmodulin revealed that the 240K polypeptide is the component that binds calmodulin. This polypeptide may be related to spectrin and fodrin, which have been reported to bind calmodulin.

Gordon *et al.*² recently reported an observation that may be related to Hsu and Molday's finding. While examining cGMP-gated cation channels in membrane patches excised from frog photoreceptors, they noticed that the channels developed a stronger affinity for cGMP during the course of their experiments. This time-dependent effect was attributed to protein dephosphorylation, as protein phosphatase inhibitors blocked

1. Ryan, M. J., Fox, J. H., Wilczynski, W. & Rand, A. S. *Nature* **343**, 66-67 (1990).
2. Basolo, A. L. *Science* **250**, 808-810 (1990).
3. Searcy, W. A. *Am. Zool.* **32**, 71-80 (1992).
4. Kirkpatrick, M. & Ryan, M. J. *Nature* **350**, 33-38 (1991).
5. Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge University Press, 1982).
6. Zahavi, A. *J. theor. Biol.* **53**, 205-214 (1975).
7. Fitzgibbon, C. D. & Fanshaw, J. H. *Behav. Ecol. Sociobiol.* **23**, 69-74 (1988).
8. Pomiankowski, A. *Proc. R. Soc. B* **247**, 123-145 (1987).
9. Grafen, A. *J. theor. Biol.* **144**, 473-516 (1990).
10. Grafen, A. *J. theor. Biol.* **144**, 517-546 (1990).
11. Johnstone, R. A. & Grafen, A. *Proc. R. Soc. B* **248**, 229-233 (1992).
12. Dumbacher, J. P. *et al. Science* **258**, 799-801 (1992).
13. Guilford, T. & Dawkins, M. S. *Evolution* (in the press).
14. Maynard Smith, J. & Harper, D. *Phil. Trans. R. Soc. B* **331**, 557-570 (1988).

the effect and application of protein phosphatase I mimicked it. Although they did not specifically address the role of calmodulin in their experiments, Gordon *et al.* found that the time-dependent change in cGMP affinity required Ca^{2+} . When EGTA was used to lower free Ca^{2+} to less than 1 nM, the channel remained in its low-affinity state. However, Hsu and Molday's finding suggests that chelating Ca^{2+} would release calmodulin from the channel complex and generate the high-affinity channel. Gordon's observation seems to be at odds with Hsu and Molday's finding, but the potentially complex relationship between calmodulin, phosphorylation state and the cGMP-dependent channel in excised patches will require closer inspection.

In another related report, Kramer and Siegelbaum³ demonstrate that Ca^{2+} affects a cAMP-gated channel in excised membrane patches derived from catfish olfactory cilia. Raising Ca^{2+} from 100 nM to 70 μM shifted the $K_{1/2}$ of this channel from 2.8 μM to 6.2 μM cAMP. This Ca^{2+} -induced decrease in affinity for cAMP appeared to depend on a soluble factor that eluted from the patches during the experiments. Might this factor be calmodulin? Kramer and Siegelbaum suggest that it isn't. Calmidazolium, a potent calmodulin inhibitor, did not block the Ca^{2+} effect and application of calmodulin to washed-out membranes did not restore Ca^{2+} sensitivity. Kramer and Siegelbaum conclude that a Ca^{2+} -binding factor other than calmodulin may mediate the effect of Ca^{2+} on the olfactory channel.

Despite their discrepancies, each of these reports provides evidence that ligand affinities of cyclic-nucleotide-gated cation channels in sensory neurons can be regulated and that Ca^{2+} participates in this regulation. What advantage might a channel with Ca^{2+} -dependent affinity for cyclic nucleotides give to a sensory neuron? Hsu and Molday propose that in photoreceptors it may promote recovery following photoexcitation. When light stimulates hydrolysis of cGMP and the concentration of intracellular Ca^{2+} falls, release of Ca^{2+} from calmodulin would strengthen the affinity of the channel for cGMP. This would promote recovery by stimulating

channels to rebind cGMP. In photoreceptors and in olfactory sensory neurons, Ca^{2+} also seems to mediate adaptation to background stimuli^{3,6}. Kramer and Siegelbaum propose that odorant-stimulated elevation of intracellular Ca^{2+} may promote adaptation by reducing the affinity of cyclic-nucleotide-gated channels for cAMP.

These findings add a new twist to the increasingly complex picture of how sensory neurons respond, recover and adapt to their stimuli. □

James B. Hurley is in the Howard Hughes Medical Institute and the Department of Biochemistry at the University of Washington, Seattle, Washington 98195, USA.

SYNTHETIC CHEMISTRY

Nature's anticancer agents

John Mann

A NEW step in the fight against cancer has been achieved with the laboratory synthesis in gram quantities of a natural compound with powerful anti-tumour activity. The synthesis, described by Nicolaou *et al.* in the *Journal of the American Chemical Society*¹, is a remarkable triumph of chemical techniques.

Despite the recent advances in our knowledge of the molecular biology of cancer induction and progression, eradication of the disease is still achieved by the classical means of surgery, radiotherapy and, most importantly, chemotherapy. In fact, chemotherapy rather than immunotherapy is likely to retain its pre-eminence for many years. In consequence, new drugs with novel modes of action are keenly sought, and the rapidly expanding research activity on the enediyne exemplifies the interest in this area.

Four natural products with an ene-

diyne subunit (compounds 1–4, Fig. 1) have been isolated from bacteria during the past 12 years^{2–5}, each possessing marked anti-tumour activity. Unlike the alkylating agents (N-mustards, platinum drugs, and so on), which damage nuclear DNA through the formation of intra-strand cross-links, these new compounds are metabolized into active diradicals

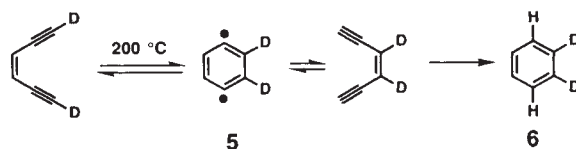


FIG. 2 The Bergman cycloaromatization reaction.

which abstract hydrogen atoms from the sugar-phosphate backbone of DNA.

The mechanism of this process can be compared with the so-called Bergman reaction which was first described in 1972⁶ (Fig. 2). An enediyne was converted into a benzene diradical (5), and this abstracts two hydrogens to produce a substituted benzene (6). The natural

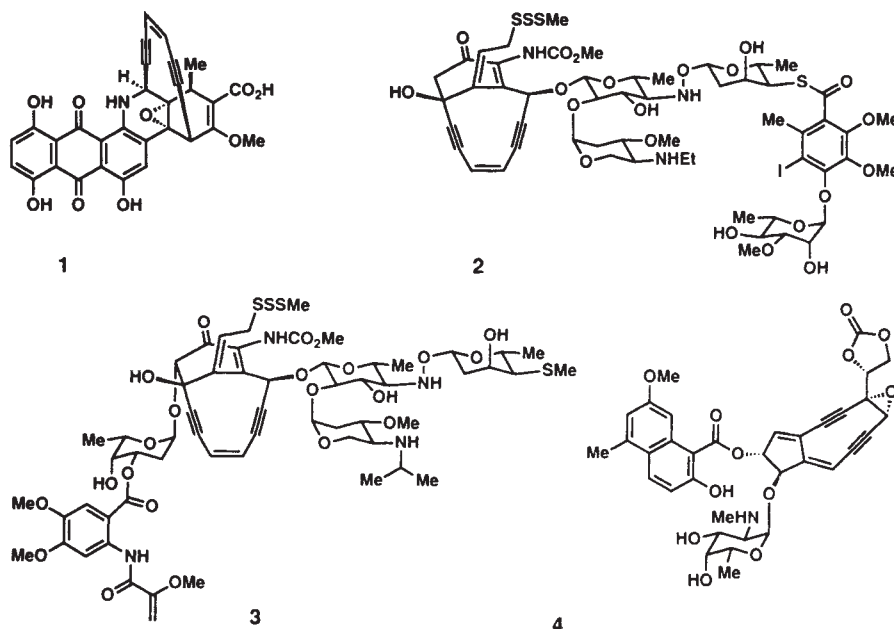


FIG. 1 Structures of naturally occurring enediyne anticancer antibiotics: 1, dynemicin A; 2, calicheamicin γ_1 ; 3, esperamicin A₁; 4, neocarzinostatin chromophore.

1. Fesenko, E. E., Kolesnikov, S. S. & Lyubarsky, A. L. *Nature* **313**, 310–313 (1985).
2. Gordon, S. E., Brautigan, D. L. & Zimmerman, A. L. *Neuron* **9**, 739–748 (1992).
3. Kramer, R. H. & Siegelbaum, S. A. *Neuron* **9**, 897–906 (1992).
4. Hsu, Y. T. & Molday, R. S. *Nature* **361**, 76–79 (1993).
5. Stryer, L. *J. Biol. Chem.* **266**, 10711–10714 (1991).
6. Yau, K.-W. *Curr. Opin. Neurobiol.* **1**, 252–257 (1991).
7. Koch, K.-W. & Stryer, L. *Nature* **334**, 64–66 (1988).
8. Kawamura, S. & Murakami, M. *Nature* **349**, 420–423 (1991).
9. Dizhoor, A. M. *et al.* *Science* **251**, 915–918 (1991).
10. Lambrecht, H. G. & Koch, K. W. *EMBO J.* **10**, 793–798 (1991).
11. Gray-Keller, M. P., Polans, A. S., Palczewski, K. & Detwiler, P. B. *Neuron* (in the press).

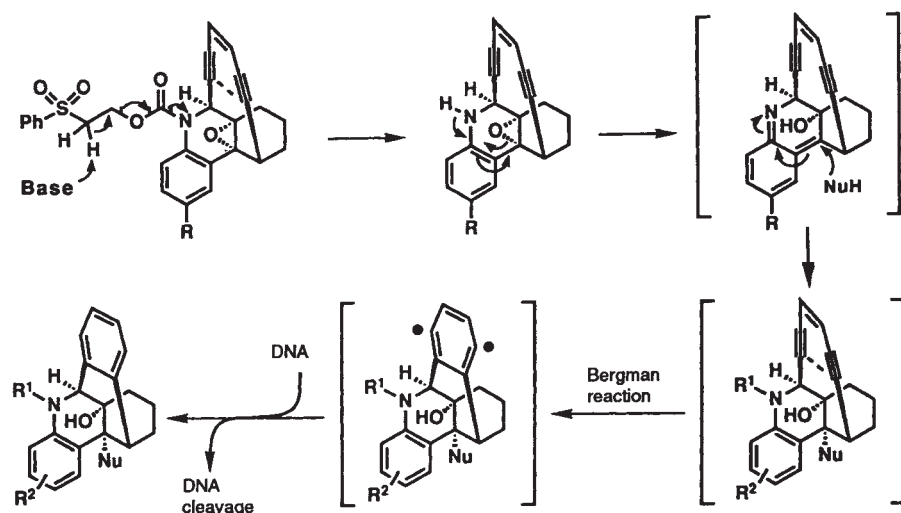


FIG. 3. Enediyne design, synthesis and biological action.

products all possess an enediyne unit as part of an overall structure which also contains a delivery system (carbohydrate or intercalating group), and a triggering device. The molecules are in effect pro-drugs, in that the two triple bonds are held apart until after the triggering mechanism has taken place. The Berg-

man reaction occurs, and migration of a proton, and the epoxide is opened thus effecting relaxation of the rigid structure, and precipitating the Bergman reaction.

This is clearly not the end of the story. Bergman could surely not have predicted that his 'innovative' chemistry had

CYTOTOXICITIES OF DESIGNED ENEDIYNE **7** AGAINST 19 TUMOUR CELL LINES (TOP) AND FOUR NORMAL CELL LINES (BOTTOM)

Cell type	Cell line	IC ₅₀ (M)	Cell type	Cell line	IC ₅₀ (M)
Melanoma	SK-Mel-28	3.1×10^{-6}	Lung carcinoma	UCLA P-3	9.8×10^{-8}
Melanoma	M-14	1.6×10^{-6}	Pancreatic carcinoma	Capan-1	3.1×10^{-9}
Melanoma	M-21	1.6×10^{-6}	T-cell leukaemia	TCAF	1.1×10^{-9}
Colon carcinoma	HT-29	1.6×10^{-6}	Multidrug resistant T-cell leukaemia	TCAF-DAX	1.7×10^{-9}
Ovarian carcinoma	Ovar-3	7.8×10^{-7}	Myeloma	RPMI-8226	7.7×10^{-9}
Astrocytoma	U-87 UG	7.8×10^{-7}	Mouse leukaemia	P-388	4.6×10^{-9}
Glioblastoma	U-251 MG	3.9×10^{-7}	Mouse leukaemia	L-1210	1.3×10^{-9}
Breast carcinoma	MCF-7	7.8×10^{-7}	Promyelocytic leukaemia	HL-60	3.6×10^{-11}
Lung carcinoma	H-358	2.0×10^{-7}	T-cell leukaemia	Molt-4	2.0×10^{-14}
Lung carcinoma	H-522	9.8×10^{-8}			
Bone marrow	HNBM	5.0×10^{-5}	Normal human dermal fibroblast	NHDF	5.0×10^{-6}
Human mammary epithelial cells	HMEC	6.3×10^{-6}	Chinese hamster ovary	CHO	3.1×10^{-6}

man reaction can then occur, and a benzene diradical is formed, leading to damage of the DNA.

A number of groups are active in this area, but Nicolaou and coworkers have probably made the greatest contributions to our knowledge of the chemistry and modes of action of these intriguing molecules^{7,8}. In their latest contribution¹, they report the first total synthesis of calicheamicin γ_1 (**2**) in stereochemically pure form. The synthesis is highly convergent and should provide access to gram quantities of the natural product and structural analogues.

In addition, Nicolaou *et al.* have also recently provided details of the synthesis of much of the dynemicin molecule, as well as of the syntheses of several simplified diynes^{9,10}. Several of these, most notably (**7**), possess quite staggering anti-tumour activity (see table), and Nicolaou proposes that the mechanism of action is as shown in Fig. 3. The enediyne is held apart in the rigid parent

structure, but following abstraction of a proton, and migration of electrons, the epoxide is opened thus effecting relaxation of the rigid structure, and precipitating the Bergman reaction.

John Mann is in the Department of Chemistry, University of Reading, Whiteknights, PO Box 224, Reading RG6 2AD, UK.

- Nicolaou, K. C. *et al.* *J. Am. chem. Soc.* **114**, 10082–10084 (1992).
- Konishi, M. *et al.* *J. Antibiot.* **42**, 1449–1452 (1989).
- Lee, M. D., Ellestad, G. A. & Borders, D. B. *Acc. chem. Res.* **24**, 235–243 (1991).
- Konishi, M. *et al.* *J. Antibiot.* **38**, 1605–1609 (1985).
- Napier, M. A., Holmquist, B., Strydom, D. J. & Goldberg, I. H. *Biochem. biophys. Res. Commun.* **89**, 635–642 (1979).
- Bergman, R. G. *Acc. chem. Res.* **6**, 25–31 (1973).
- Nicolaou, K. C., Dai, W.-M., Tsay, S.-C., Estevez, V. A. & Wrasido, W. *Science* **256**, 1172–1178 (1992).
- Nicolaou, K. C. & Smith, A. L. *Acc. chem. Res.* **25**, 497–503 (1992).
- Nicolaou, K. C. *et al.* *J. Am. chem. Soc.* **114**, 8890–8907 (1992).
- Nicolaou, K. C. & Smith, A. L. *J. Am. chem. Soc.* **114**, 8908–8921 (1992).

DAEDALUS

That inward eye

SOME forms of brain surgery are conducted under purely local anaesthetic. When the surgeon touches the conscious brain with an electrode, the patient is often overwhelmed by some vivid memory. Different regions of the brain trigger different memories. Daedalus now proposes a less invasive method of exploring our mental world.

He points out that a changing magnetic field can induce a strong but diffuse ring current in an electrical conductor. To give a more concentrated ring current, he has devised an electromagnet whose two pole pieces are tubular. The field between them is hollow, less intense in the middle than at the periphery. Increase the current in the electromagnet, and the strong annular field expands into the central region. A conductor in that region experiences a small, localized ring current. Two such magnets at right angles, their fields intersecting in quadrupolar fashion, define an even tighter ring current at the point of intersection.

Centred on a human head, and given a suitable sudden step-increase in its current, Daedalus's quadrupolar magnet should induce a ring current in a small specific area of the brain. The local neurons would be fired. The subject would carry out whatever action, experience whatever sensation, or recall whatever memory was triggered by those neurons.

DREADCO volunteers are now trying this out in practice. For safety's sake, they operate the whole machine themselves. They choose the region of their brain to be activated, and slowly increase the intensity of the current pulses of the magnet until they start to feel the effects. If these are too disconcerting, they can back off or move to another region; if the effects are interesting or pleasant they can increase the intensity further. Each volunteer can thus explore his brain, and produce a crude 'map' of its contents.

The results will be fascinating, particularly for Daedalus's more absent-minded volunteers. Will they find all sorts of mislaid memories, quite beyond conscious recall? Does the brain, like some obsessive bureaucracy, really file copies of everything? Do some people pack a lot more memory into a given space, by efficient coding? If so, what is the code? Daedalus suspects that information on a particular topic (like animals or faces) may well be coded in the form of an archetype, with specific examples being economically stored as perturbations of that archetype. The DREADCO encephalomagnet may reveal at last the primitive elements from which we construct our image of the world.

David Jones

Marine litter keeps increasing

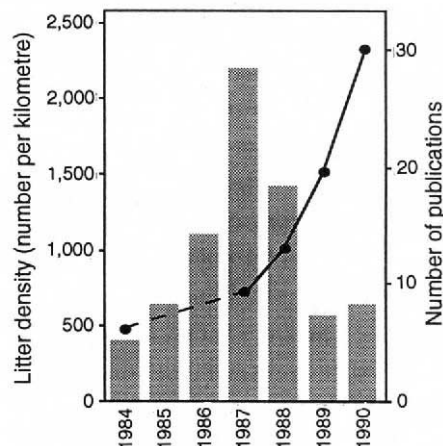
SIR — In addition to being aesthetically unappealing, litter has direct impacts on marine ecosystems. It provides novel attachment sites for sessile organisms, kills or injures animals that become entangled in or eat litter, and possibly aids dispersal of terrestrial organisms¹. The amount of litter at sea is increasing despite control measures.

Floating litter drifts throughout the world's oceans, either dumped from ships or blown and washed from land. Even Antarctica, far from source areas, is affected². Plastic articles pose the greatest problem: once at sea, plastics are virtually immune to degradation and can drift for years, covering vast distances. Steps taken to curb the release of persistent litter into the sea include education, product substitution and legislation. The most far-reaching step was taken in 1987, when countries responsible for more than half the world's shipping acceded to Annex V of the International Convention for the Prevention of Pollution from Ships (MARPOL), which bans the disposal of persistent wastes at sea¹. This measure came into force at the end of 1988, but there have been few attempts to monitor its effectiveness.

Between 1984 and 1990, we surveyed litter washed up at Inaccessible Island (37°15' S, 12°30' W), a remote island in the Tristan da Cunha group, south Atlantic Ocean^{3,4}. The island is uninhabited and is seldom visited, making it ideal for monitoring the rate of litter accumulation. The figure shows how the density of litter increased exponentially throughout the study period, with MARPOL Annex V having no discernible impact on the rate of increase. Plastic articles comprised more than 80% of the litter, with most identifiable items coming from South America, more than 3,000 km distant.

The exponential increase in marine litter throughout the latter half of the 1980s contrasts with the amount of scientific research on the subject. Numbers of publications addressing the marine litter problem peaked in 1987 (see figure), coinciding with the decision bringing into force MARPOL Annex V. The recent decrease in publications may stem from a perception in some quarters that the threat posed by marine litter has been identified adequately, and that resources are better allocated to control measures than to research and monitoring. This viewpoint is simplistic. The impact of marine litter remains poorly understood. We know little about the effects of entanglement and ingestion on marine populations, and there has been no study of the consequences of greatly increasing habitat availability for sessile organisms

in oceanic ecosystems. Until these questions are addressed, we cannot rank marine litter relative to other environmental problems, and thus cannot sensibly decide how much attention it war-



ants. Further research is required both to assess the impact of marine litter and to sustain efforts to curb its release.

Peter G. Ryan
Coleen L. Moloney
 FitzPatrick Institute,
 University of Cape Town,
 Rondebosch 7700,
 South Africa

- Shomura, R. S. & Godfrey, M. L. (eds) *Proc. 2nd Int. Conf. Marine Debris*, NOAA Tech. Mem. NMFS **154**, 1–1274 (1990).
- Gregory, M. R., Kirk, R. M. & Mabin, M. C. G. N. Z. *Antarct. Rec.* **7** (3), 32–47 (1984).
- Ryan, P. G. *Envir. Conserv.* **14**, 341–346 (1987).
- Ryan, P. G. *Proc. 2nd Int. Conf. Marine Debris*, NOAA Tech. Mem. NMFS **154**, 85–102 (1990).

Ribosomal RNA trees misleading?

SIR — Phylogenetic placing of protozoa that lack mitochondria is crucial in clarifying the early evolution of eukaryotes¹. Most studies relevant to this problem have been done using small-subunit ribosomal RNA (ss rRNA) sequences^{2–4}. These studies suggest that, among three protozoa lacking mitochondria, *Giardia lamblia* and *Vairimorpha necatrix* represent the earliest and the second earliest offshoots in the eukaryotic tree, whereas *Entamoeba histolytica* separated from higher eukaryotes after *Euglena gracilis* and *Trypanosoma brucei* (both with

mitochondria) diverged. However, we suggested^{5–7} that this tree could be erroneous because G+C content of ss rRNA differs drastically among species, and because this effect has not been taken into account in inferring the ss rRNA tree. For example, the G+C content of *G. lamblia* is as high as 74.7%, while those of *V. necatrix* and *E. histolytica* are as low as 37.7 and 38.3%.

From a maximum likelihood^{8,9} analysis of amino-acid sequence data of elongation factor 1 α , we have now shown⁵ that *E. histolytica* is likely to be an outgroup to *E. gracilis* and higher eukaryotes, in disagreement with the ss rRNA tree. Furthermore, from a parsimony analysis of amino-acid sequence data of DNA-dependent RNA polymerase III, Lanzendorf *et al.*¹⁰ suggested that *G. lamblia* may not be an outgroup to *T. brucei* and higher eukaryotes, again in disagreement with the ss rRNA tree. Although statistically inconclusive, we reconfirmed⁶ their suggestion by using the maximum likelihood method.

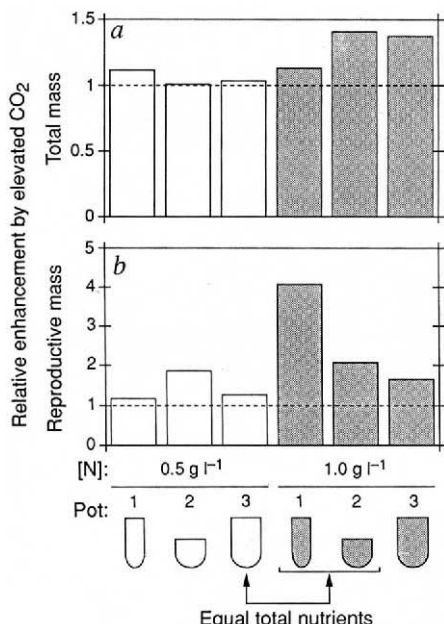
We have further shown that the amino-acid compositions of these highly conservative proteins are relatively free from the variation of biased base compositions of DNA and ss rRNAs (refs 6,7). From these results, we suggest that the ss rRNA tree representing early branchings of eukaryotes must be re-examined by using amino-acid sequences of conservative proteins such as elongation factors, RNA polymerases and ATPases. The ss rRNA sequences have been widely used in clarifying deep branchings in phylogenetics, and they have provided us with invaluable information. Nevertheless, we must be careful in using these data, particularly when the G+C content differs among species¹¹.

Masami Hasegawa
Tetsuo Hashimoto
 Institute of Statistical Mathematics,
 4-6-7 Minami-Azabu,
 Minato-ku,
 Tokyo 106,
 Japan

- Cavalier-Smith, T. *Nature* **339**, 100–101 (1989).
- Vossbrinck, C. R. *et al.* *Nature* **326**, 411–414 (1987).
- Sogin, M. L. *et al.* *Science* **243**, 75–77 (1989).
- Sogin, M. L. *Curr. Opin. Genet. Dev.* **1**, 457–463 (1991).
- Hasegawa, M., Hashimoto, T., Adachi, J., Iwabe, N. & Miyata, T. *J. molec. Evol.* (in the press).
- Hashimoto, T., Adachi, J. & Hasegawa, M. *Endocytobiosis Cell Res.* (in the press).
- Hasegawa, M., Hashimoto, T. & Adachi, J. in *The Origin and Evolution of Prokaryotic and Eukaryotic Cell* (ed. Hartman, H. & Matsuno, K.) (World Sci. Publ., in the press).
- Kishino, H., Miyata, T. & Hasegawa, M. *J. molec. Evol.* **31**, 151–160 (1990).
- Adachi, J. & Hasegawa, M. *MOLPHY: Programs for Molecular Phylogenetics, I. — PROTML: Maximum Likelihood Inference of Protein Phylogeny* (Institute of Statistical Mathematics, Tokyo, 1992).
- Lanzendorf, M. *et al.* *Nucleic Acids Res.* **20**, 1145 (1992).
- Loomis, W. F. & Smith, D. W. *Proc. natn. Acad. Sci. U.S.A.* **87**, 9093–9097 (1990).

Plant responses to carbon dioxide

SIR — There have been several reports suggesting that non-field-based studies of plant responses to elevated atmospheric CO₂ may be compromised because small rooting volumes limit a plant's ability to respond to elevated CO₂ (refs 1–3). This hypothesis calls into question most of our present knowledge of vegetation responses to elevated CO₂, and is not supported by new evidence (K. D. M.



CO₂-induced growth enhancement of *Abutilon theophrasti* Medic. (elevated/ambient dry mass) as a function of pot size, shape and nutrient concentration. *a*, Total dry mass enhancement before fruit maturation; *b*, reproductive dry mass enhancement immediately after fruit maturation. Pots were 0.65 l (tall and narrow & short and wide pots) or 1.30 l in volume; nutrient concentrations were 0.5 g l⁻¹ or 1.0 g l⁻¹ of a balanced N-P-K fertilizer added to each pot every 3 days in sufficient amount to displace the soil solution present. Large pots of low nutrient concentration and small pots of high nutrient concentration received equal total nutrients (K. D. M. McC., G. M. B. and F. A. B., manuscript submitted).

McC., G. M. B. and F. A. B., manuscript submitted).

Elevated CO₂ atmospheres stimulate photosynthesis (and growth) when photosynthetic products (sugars and starches) are used at a sufficient rate⁴. Plants grown in small pots have restricted root growth and thus reduced photosynthate demand^{1,3}. However, a reduction in photosynthesis is not a necessary outcome of root restriction; photosynthetic rates may increase or not change.

Arp¹ reviewed published values of CO₂-induced photosynthetic enhance-

ment, concluding that non-field-based studies suffer from root restriction and are thus unreliable tests of vegetation responses in the field. However, he did not examine correlations between the photosynthetic rates and aspects of nutrient supply, which can decrease with pot size.

In one of only two studies that directly examined the effects of pot size on CO₂ responsiveness, cotton (*Gossypium hirsutum*) grown in small pots had less CO₂-induced photosynthetic enhancement³. CO₂-induced growth enhancement was greatest in small pots, however. Leaf starch and sucrose contents were unaffected by pot size but were greater for high CO₂-grown plants in both small and large pots. The authors did not examine whether reduced nutrient supply, a function of soil volume, may have caused or contributed to the reduced photosynthesis.

We have examined the influence of both physical rooting space and soil nutrients (K. D. M. McC., G. M. B. and F. A. B., manuscript submitted). CO₂-induced growth and reproductive output enhancement are greatest in pots with greater nutrient concentrations, regardless of pot volume or total nutrient content (see figure). Enhancement of reproductive output is greater in plants growing in smaller volumes, particularly when the nutrient concentration within the pot was high. Remarkably, pot shape *per se* also influences growth and repro-

ductive enhancements.

These results bear on the potential limitations to CO₂ enhancement in pot studies in two ways. First, small pots do not necessarily limit CO₂ responsiveness; they may not affect it or they may increase it. Second, CO₂ responsiveness is dependent on many aspects of growth conditions within pots, especially nutrient concentration and the physical dimensions of the pot.

The suggestion that small pots limit the CO₂ responsiveness of plants is compelling, but not supported, and it is premature to abandon decades of non-field-based research. Furthermore, plants in the field do not have unlimited below-ground resources with which to maximize growth in a CO₂-rich world. It would be risky to devise policies based on the presumption that the CO₂ fertilization effect will be sustained indefinitely in natural systems.

K. D. M. McConaughay*

G. M. Berntson

F. A. Bazzaz

Department of Organismic and Evolutionary Biology,
Harvard University,
Cambridge, Massachusetts 02138, USA

*Permanent address and address for correspondence: Department of Biology, Bradley University, Peoria, Illinois 61625, USA.

1. Arp, W. J. *Pl. Cell Envir.* **14**, 869–875 (1991).
2. Idso, S. B. *Tellus* **43B**, 338–341 (1991).
3. Thomas, R. B. & Strain, B. R. *Pl. Physiol.* **96**, 627–634 (1991).
4. Stitt, M. *Pl. Cell Envir.* **14**, 741–762 (1991).

A palatal rete in the right whale?

SIR — It has recently been suggested by Ford and Kraus¹ that the palate of the right whale (*Eubalaena glacialis*) possesses a previously undescribed vascular rete that has a thermoregulatory function. A rete is a specialized complex network of fine arteries and/or veins. Ford and Kraus suggest that right whales (Balaenidae) can dump heat by opening the mouth and allowing a flow of cold water over the hard palate. We find several potential flaws with this interpretation.

The right whale is not unique in having a highly vascularized palate. The palate of most mammals has a rich blood supply including arterial anastomoses and a venous plexus^{2–4}. In all baleen whales the palate contains numerous small blood vessels and nerves associated with the formation of baleen that these whales use for filter feeding. This observation in rorqual whales (Balaenopteridae) previously led to the speculation that the palate may function in thermoregulation⁵. Our gross observations of the right whale suggested only a typical baleen whale vascular palate, not a unique rete system. We find no mor-

phological evidence in Ford and Kraus's report that right whales possess a true specialized rete. It is also not clear how mechanoreceptors function directly in thermoregulation as proposed by these authors. Mechanoreceptors are probably involved in detecting water flow through the mouth and over the baleen as these whales feed.

Several additional aspects of the morphology and behaviour of right whales do not support the hypothesis that the palate functions to dump body heat. The palate of right whale is long, but extremely narrow. Therefore, the surface area is relatively small, especially when compared to other baleen whales. In addition, the palatal surface would be exposed to a flow of water only when the mouth is open. This occurs as right whales engage in skim feeding, but this species primarily feeds in higher lati-

1. Ford, T. J. & Kraus, S. D. *Nature* **359**, 680 (1992).
2. Evans, H. E. & Christensen, G. C. *Miller's Anatomy of the Dog* (Saunders, Philadelphia, 1979).
3. Crouch, J. E. *Text-Atlas of Cat Anatomy* (Lea and Febiger, Philadelphia, 1969).
4. Brash, J. C. & Jamieson, E. B. (eds) *Cunningham's Text-Book of Anatomy* (Oxford University Press, New York, 1937).
5. Slijper, E. J. *Whales* (Basic, New York, 1962).

tudes and breeds in lower latitudes. Therefore they only exhibit a mouth-open behaviour when they are in colder waters, not in warmer waters when they would be more likely to be thermally stressed. More work is needed to document all the potential functions of the baleen whale palate.

John E. Heyning

Section of Birds and Mammals,
Natural History Museum of
Los Angeles County
900 Exposition Blvd, Los Angeles,
California 90007, USA

James G. Mead

Division of Mammals,
National Museum of Natural History,
Washington, DC 20560, USA

Michael M. Bryden

Department of Veterinary Anatomy,
University of Sydney,
New South Wales, 2006, Australia

Climate and cocaine

SIR — It is difficult to obtain information about the production of clandestine crops such as the coca (*Erythroxylum coca*) leaf. We have used production of coca leaf in Bolivia as a test of the effectiveness of remote climate analysis in obtaining such surveillance information. Bolivian coca production provides about 40% of the world's supply of illicit cocaine¹ but is difficult to monitor be-

cause of the severe restrictions on its sale.

We studied coca-leaf production influenced by fluctuations in rainfall at distant Bolivian stations. Monthly data (1976–81) of production for one of the two coca-producing regions of Bolivia, the Chapare (see figure), were correlated with the number of days of rain for 15 Bolivian stations. We restricted study to the high-Sun rainy season (November–February) which comprises 40–50% of the annual coca production period¹. Our monthly coca production values were obtained from Bolivian government records compiled from the checkpoint on the single road to the Chapare region² and were standardized by month.

Climate was represented by (1) the total monthly rainfall and (2) the number of days with rainfall 1 mm or greater. As with the coca values, our climate data were standardized by month (1961–86) as documented in ref. 3 and from locally obtained data.

Although Bolivian monthly rainfall is significantly linked to coca production, the number of days of rain per month is more closely related to Chapare coca production (see table). Almost all major Bolivian stations demon-

strate a strong inverse relationship between days of rain and coca production. Production decreases as the number of rain days increases, and this relationship is highly significant at seven locations. For example, the number of rain days at Rurrenabaque (approximately 150 km from the Chapare) explains 64% of the variance in coca production.

The relationship is not likely to be related to physiological growth factors. We propose that the number of rain days constrains the harvest by decreasing the time available to dry the picked coca leaves. In general, a two- to three-day drying time is needed to ensure a good harvest⁴. The greater the number of rain days, the fewer leaves are produced for market.

The effectiveness of this remote climate relationship between the number of

COCA PRODUCTION AND CLIMATE

Weather station	Geographical land type	Rainfall	Rain days
Camiri	Transitional	-0.59(20)†	-0.47(20)*
Charana	Altiplano	0.27(9)	-0.70(9)*
Cobija	Tropical lowland	-0.35(20)	-0.36(20)
Cochabamba	Altiplano	-0.37(20)	-0.46(20)*
Concepcion	Tropical lowland	-0.47(9)	-0.27(9)
La Paz	Altiplano	-0.55(20)*	-0.58(20)†
La Jota	Transitional	-0.68(14)†	-0.77(14)†
Oruro	Altiplano	-0.70(20)†	-0.54(20)*
Riberalta	Tropical lowland	0.30(20)	0.06(20)
Rurrenabaque	Tropical lowland	-0.03(19)	-0.80(19)†
Santa Cruz	Tropical lowland	-0.16(20)	-0.72(20)†
Sucre	Altiplano	-0.57(20)†	-0.70(20)†
Tarija	Transitional	-0.41(20)	-0.46(20)*
Trinidad	Tropical lowland	-0.02(20)	-0.63(20)†
Yacuiba	Transitional	-0.27(20)	-0.75(20)†

Correlation coefficients between high-Sun coca production in the Chapare region and two climate variables recorded at 15 Bolivian cities. Numbers in parentheses refer to the number of available months used to compute the coefficient. Asterisk, coefficient significant at $\alpha = 0.05$; dagger, coefficient significant at $\alpha = 0.01$.

rain days at various Bolivian cities and coca production is explained by the geography of the country. Cities where rain day totals are strongly linked to Chapare coca production are either geographically close to the coca-producing region (for example, La Jota, Trinidad); located in the climate region defined by the large transition zone between the lowland rainforests and the Altiplano region (for example, Santa Cruz, Rurrenabaque); or derive their precipitation from the same regional moisture flow as the coca-producing area (for example, La Paz, Sucre).

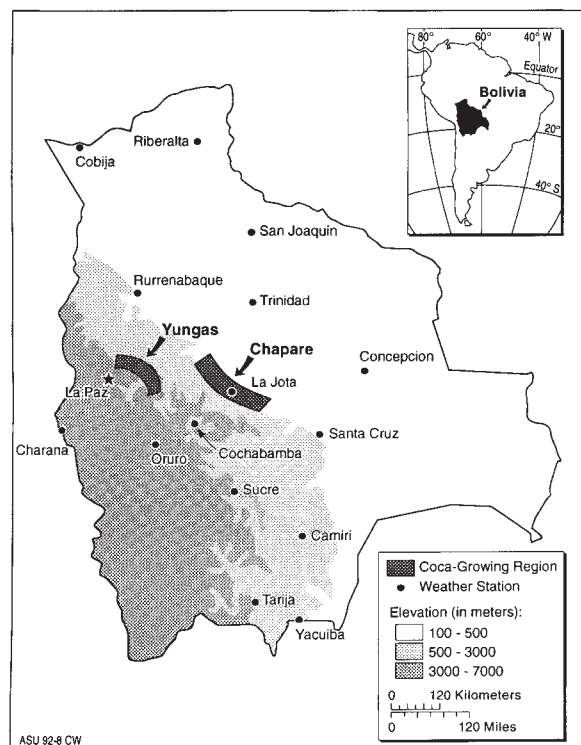
Unfortunately, the lack of recent coca harvest data¹ and on-site climate data prevents independent prediction of coca production. However, because of the strong inverse relationship described above, first-order approximations of coca production can be achieved by remote weather monitoring. In particular, the use of modern technology (such as satellite imagery) might substitute for the scarcity of on-site climate information for such sensitive regions. We suggest that the study of climate in this way may improve the monitoring, prediction and perhaps eventual control of the coca supply.

Michael S. McGlade

Randall S. Cerveny

Ray Henkel

Department of Geography,
Arizona State University,
Tempe,
Arizona 85287-0104, USA



Meteorological stations for this study, and the two primary coca-producing regions (the Chapare and Yungas regions) of Bolivia.

1. Narcotics Intelligence Consumers Committee *The Supply of Illicit Drugs to the United States from Foreign and Domestic Sources in 1988–1989* (1990).
2. Orellana, A. C. & Zannier, J. C. C. *Bolivia: Coca Cocaina (Subdesarrollo y Poder Político)* (Los Amigos Del Libro, Cochabamba, Bolivia, 1982).
3. World Meteorological Organization *Monthly Climatic Data for the World Vols 15–39* (NOAA, Asheville, North Carolina, 1962–1986).
4. Henkel, R. *The Chapare of Bolivia: a Study of Tropical Agriculture in Transition* (University of Wisconsin, 1971).

Two-dimensional crystallization

SIR — Many people have obtained two-dimensional (2D) colloid crystals on substrates¹⁻⁴, but they have observed the final results of ordering and have not investigated the mechanism of this process. We have directly observed the dynamics of 2D array formation of latex particles on solid substrate by means of optical microscopy. Our observations suggest a two-stage mechanism of 2D crystallization: (1) Nucleus formation, governed by attractive capillary forces appearing between particles partially immersed in a liquid layer⁵; and (2) Crystal growth, through convective particle flux caused by the water evaporation from the already ordered array.

A fast and convenient method for formation of a 2D protein array on a mercury surface has been recently developed^{2,3}. The good quality of the samples thus obtained allowed investigation of the protein orientation and structure by electron microscopy combined with image reconstruction. Highly ordered 2D crystals can thus be obtained in a controllable way; this technique could lead to further development of the

controlled building up of well-ordered protein monolayers and multilayers — a possible step towards a future high technology at the macromolecular level³. 2D arrays on solid substrates can find applications in some modern techniques, such as data storage, optical devices and microelectronics⁴.

In our model experiments, we investigated the mechanism of 2D crystallization using a suspension containing monodisperse particles of 1.70 μm diameter.

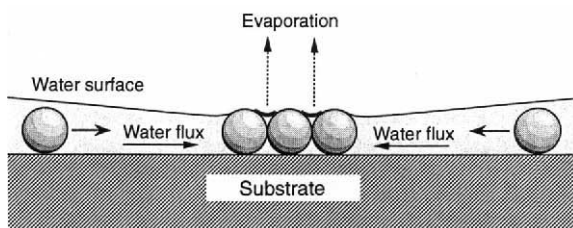


FIG. 2 Schematic presentation of the particle assembly process driven by the liquid flow.

This was spread over a horizontal hydrophilic glass plate encircled by a teflon ring. The slightly concave layer formed gradually thins owing to the water evaporation; when its thickness in the centre of the substrate becomes equal to the particle diameter, a nucleus of 2D crystal suddenly forms. The particles in the thicker layer encircling the nucleus begin to move towards the ordered zone and upon reaching the boundary of the array they are trapped in it (Fig. 1a). In some experiments we added 0.2 wt% glucose to the suspension and the flux of particles became slower (Fig. 1b).

The nature of the forces governing the ordering is revealed by the fact that in all experiments the 2D crystallization always started when the thickness of the water layer became equal to the particle diameter. This implies that the 2D-crystal nuclei are formed under the capillary attraction arising when the tops of the particles protrude from the water layer (Fig. 2). The attraction energy can be much larger than the thermal energy (kT) even with nanometre-sized particles⁵.

We were able to show that the crystal growth is caused by a convective transport of particles towards the ordered nucleus. This effect appears when menisci (shown bold in Fig. 2) form around the protruding tops of the hydrophilic particles in the nucleus. These menisci hinder the further thinning of the water layer in the nucleus. An intensive water influx from the thicker parts of the layer,

which tends to compensate the water evaporation from the nucleus, appears next. This flux carries the suspended particles towards the nucleus. By decreasing or increasing the water evaporation rate we could speed up or slow down the convective particle transport. At increased humidity, we saw a complete arrest of the process of ordering and even disintegration of the already ordered clusters.

N. D. Denkov

O. D. Velev

P. A. Kralchevsky

I. B. Ivanov*

Faculty of Chemistry,
University of Sofia,
1126 Sofia,
Bulgaria

H. Yoshimura

K. Nagayama

Protein Array Project, ERATO,
JRDC, 18-1 Higashiara,
Tsukuba,
305 Japan

* To whom correspondence should be addressed.

Checkpoint check

SIR — Andrew Murray in his interesting and comprehensive Review Article (*Nature* **359**, 599–604; 1992) misnames the fission-yeast checkpoint genes that we have identified (Fig. 3). Our paper that is cited in support of the information in fact reported mutants rather than genes. These mutants define five new checkpoint genes called *hus1-hus5*. Mutations of any one of the genes prevent arrest in response to inhibition of DNA replication and in addition cause increased sensitivity to radiation. Thus it is unlikely that any one of them is involved solely in detection of unrepliated DNA as Murray's Fig. 3 suggests. A full report on our work has appeared in *Genes and Development* (**6**, 2035–2046; 1992).

Tamar Enoch

Anthony Carr

Paul Nurse

ICRF Cell Cycle Group,
Department of Biochemistry,
University of Oxford,
Oxford OX1 3QU, UK

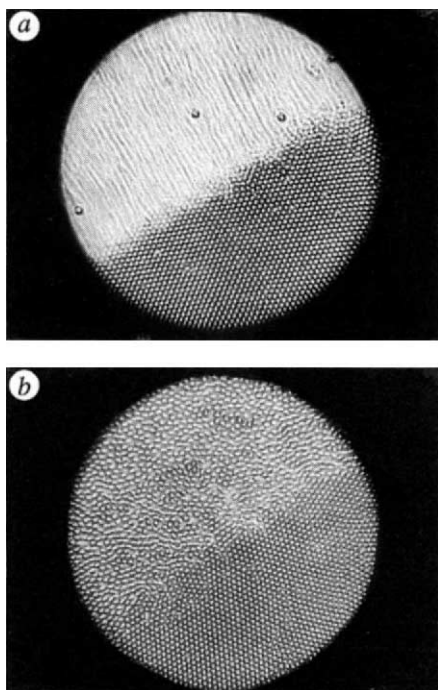


FIG. 1 Photographs of 2D-crystal growth: a, tracks of the particles rushing towards the ordered phase; b, at decreased rate of water evaporation the velocity of the moving particles is lower and the tracks are shorter.

1. Alfrey, T. Jr, Bradford, E. B. & Vanderhoff, J. W. *J. Opt. Soc. Am.* **44**, 603–609 (1954).
2. Yoshimura, H., Matsumoto, M., Endo, S. & Nagayama, K. *Ultramicroscopy* **32**, 265–274 (1990).
3. Nagayama, K. *Nanobiology* **1**, 25–37 (1992).
4. Hayashi, S., Kumamoto, Y., Suzuki, T. & Hirai, T. *J. Colloid Interface Sci.* **144**, 538–547 (1991).
5. Kralchevsky, P. A., Paunov, V. N., Ivanov, I. B. & Nagayama, K. *J. Colloid Interface Sci.* **151**, 79–94 (1992).

Scientific Correspondence

Scientific Correspondence is a relatively informal section of *Nature* in which matters of general scientific interest, not necessarily those arising from papers appearing in *Nature*, are published. Because there is space to print only a small proportion of the letters received, priority is usually given according to general interest and topicality, to contributions of fewer than 500 words, and to contributions using simple language.

Digital daughter

Jack Meadows

Ada, the Enchantress of Numbers: A Selection from the Letters of Lord Byron's Daughter and Her Description of the First Computer. Edited by Betty Alexandra Toole. Strawberry (Critical Connection, PO Box 452, Sausalito, CA 94966, USA): 1992. Pp. 439. \$29.95, £24.99. (Distributed in the UK by Pickering and Chatto.)

CHILDE Harold's *Pilgrimage* was the poem that made Byron famous. His continuation of it in 1816 began with a reference to his daughter, Ada, born the year before:

Is thy face like thy mother's, my fair child!
Ada! sole daughter of my house and heart?

Byron was not to learn much about his daughter. His wife, whom he characterized as the "Mathematical Medea", separated from him, and he departed to Italy. Lady Byron determined that their daughter should follow her interests, not his — to the extent that William Frend, who was Lady Byron's tutor and became Ada's, warned her against overpressurizing the child. Frend was well-acquainted with the leading English mathematicians, so it is not surprising that Ada in her teens encountered Charles Babbage. It is for her friendship with him that she is remembered by historians of science.

Ada met Babbage in 1833, at a time when his difference engine was well-advanced, and she was soon invited to inspect it. Her interest in Babbage then waned for some years while she married and had a family. In the early 1840s, by which time she had become Countess of Lovelace, it waxed again. Babbage had by now moved on from the difference engine to the analytical engine. He encouraged Ada to translate an article about it from the French and to add her own extensive notes. The result has become an important source of information on Babbage's computer. It was the high point of Ada's contribution. Her health subsequently deteriorated and she died at an early age in 1852.

How historically important was Lady Lovelace? More especially, how scientifically important was she? Does she merit a volume of collected letters? She was certainly intelligent and the possessor of a lively mind. But reaching a just estimate of her abilities is made difficult by her own very high opinion of herself. For example, she reports to her mother that: "Faraday expresses himself in absolute amazement at what he (I think most happily and beautifully) designates the 'elasticity of my intellect'." Although she then comments on Faraday's modesty, she concludes: "I may be the Deborah, the Elijah of Science".

The same faith in her own abilities shines through her correspondence with Babbage about the article on the analytical engine: "I wish you were as accurate, and as much to be relied on, as I am myself. . . . I hope you do not take upon yourself to alter any of my corrections". This seems a rather cavalier way of writing to an eminent scientist, 25 years her senior, about one of his inventions. The surprise is not that the irascible Babbage finally exploded, but that it took him so long to do so. Ada thought all the blame was his: "I am sorry to have to come to the conclusion that he is one of the most impracticable, selfish, and intemperate persons one can have to do with. . . . He was furious; I imperturbable and unmoved".

It is difficult to obtain a balanced view of Ada from the selection in this book, because it includes only letters from her, not to her. The editing is also sometimes less than helpful. For example, it seems rather odd to be told: "I have deleted all mathematical formulas, except in Ada's notes, because Ada's mathematical correspondence represents what she did not know, not what she knew". Doubtless

the omissions make little difference, but they reflect the defensive attitude of the editor. She clearly wishes to press the claims of Lady Lovelace as a significant thinker. This can lead to some rather Whiggish history, as, for example, in the explanations as to why it is especially appropriate for Ada to have a programming language named after her. More to the point, there is little evidence from Ada's letters that she would have successfully pioneered new developments in science had she lived longer. Her ideas on what she hoped to do were set down in a letter written not long before her final illness removed her motivation:

I have my hopes, and very distinct ones too, of one day getting cerebral phenomena such that I can put them into mathematical equations; in short a law or laws, for the mutual actions of the molecules of brain; (equivalent to the law of gravitation for the planetary and sidereal world).

Scientific knowledge of the brain remained at a low level throughout the nineteenth century. General laws, certainly laws on a par with those deduced by Newton, would have been well beyond the scope of a nineteenth-century scientist. Ada, had she continued her work, must either have changed direction or failed to make much progress.

This leads back to the original question of Ada's historical status. In terms of what she did in science, Lady Lovelace hardly ranks on a par with her contemporary, Mrs Somerville (whom she knew and admired). Nor do her



Flights of fancy — the short-winged grasshopper *Calliptamus* (left) and the spider-hunting wasp *Sceliphron destillatorium*, taken from John Brackenbury's *Insects in Flight* (Cassell, £18.99). This informative book contains over 200 of the author's striking colour photographs, which reveal the intricate and diverse techniques of flight used by moths, butterflies and other insects.

letters suggest that, given a few years more, she would have contributed to new, major breakthroughs. Of the letters selected for publication here, no more than a quarter will interest historians of science, and they add little to the picture painted some 15 years ago by Doris Moore in her biography of Lady Lovelace. Where the interest may lie is in the light that the letters throw on the problems of Byron's family — a conclusion that would have thoroughly annoyed Ada herself. □

Jack Meadows is in the Department of Information and Library Studies, Loughborough University, Loughborough, Leicestershire LE11 3TU, UK.

Carl who?

Tim Lincoln

The Pill, Pygmy Chimps, and Degas' Horse: The Remarkable Autobiography of the Award-Winning Scientist Who Synthesized the Birth Control Pill. By Carl Djerassi. *Basic Books: 1992. Pp. 319. \$25.*

ALONGSIDE the me-too title and the subtitle (half an essay in itself), a stern Carl Djerassi looks out from the dust jacket of his autobiography. On the back, Stephen Jay Gould all but calls Djerassi a Renaissance man, and Linus Pauling says he neglected his work for a couple of days because he found the book so compelling. Who can disagree? Djerassi has lived a life well worth the telling. Here he does so deftly, throwing in some soul searching about the dark corners of his personal life and commentary on science, universities and the American Way in the process.

The title is an echo of those favoured by Gould in his popular works. It is intended to hint at Djerassi's impressive range of activities and achievements. 'Pygmy chimps' reflects his involvement in a quest for a good animal model in biomedical research (which met with mixed fortunes), and 'Degas' horse' his growing interest in the arts, in later life, as poet, novelist and patron. On this second count he writes of his collection of Paul Klees and of his establishment of the Djerassi Foundation, an artist's colony, and mentions in passing his fondness for opera and (with becoming modesty) his splendid, racy novel *Cantor's Dilemma*. He also lets on that he has an aversion to paying import duty.

These stories, and the unexceptional traveller's tales, are all well and good. But it is Djerassi's incarnations as an organic chemist with entrepreneurial leanings and as a Jewish immigrant to

the United States from Europe that give his book its especial attraction and claim on a wide audience.

In 1939, aged 16, Djerassi stepped on a liner at Genoa, saying goodbye to a Vienna overrun by the Nazis and then hello to the United States. The two seem to have been made for each other. The United States offered the young Djerassi an education in chemistry and unlimited opportunities; Djerassi offered energy, ambition and a certain refusal to be constrained by norms. The upshot — told in the best two chapters, cleverly part didactic on steroid chemistry and part thriller — was the synthesis first of cortisone then of 19-nor-17 α -ethynyltestosterone (norethindrone), a highly active inhibitor of ovulation when taken orally. That was the birth of the Pill, at Syntex in Mexico City on 15 October 1951, and rumination on the consequences for Djerassi, not to mention for women, takes up much of the rest of the book.

This is mostly sobering stuff. At the centre of the book is horror at the decline of pharmaceutical research into control of fertility and contained fury over the issue of abortion. And there are, for instance, the differences of opinion with women's groups (with which, it has to be said, Djerassi has considerable empathy) and with regulatory bodies such as the Food and Drug Administration and the Environmental Protection Agency (EPA). The following quotation comes in a different context, the testing of a synthetic insect-growth regulator for mosquito control. But it gives a taste of research later in Djerassi's career, and I find it irresistible:

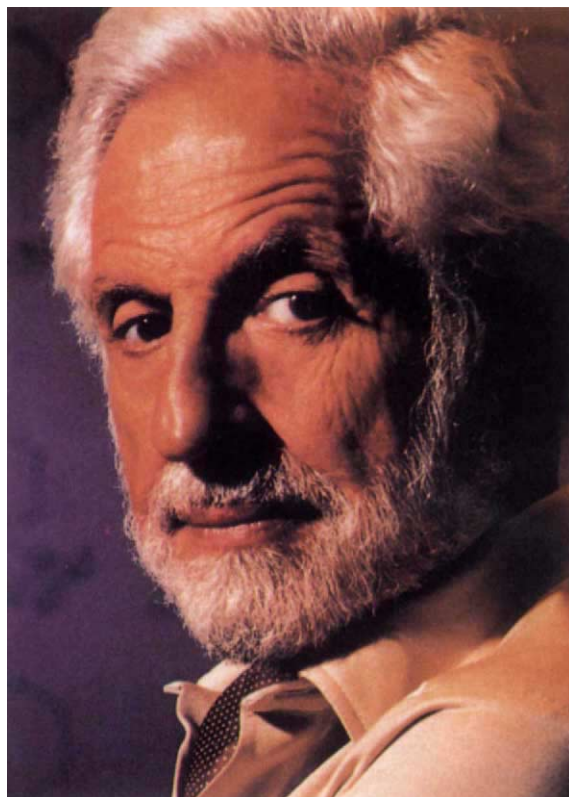
... by the time we finally received EPA approval for field applications, ALTO-SID's lack of toxicity had been demonstrated in water fleas, protozoa, copepods, sideswimmers, aquatic earthworms, mudworms, leeches, tadpoles, and snails; in mosquito fish, bluegill, trout, channel catfish, cono salmon, carp, and stickleback; in crustaceans like seashore crabs, blue crabs, mud crabs, crayfish, acorn barnacles, and various species of shrimp. By that time we didn't have to be asked about oysters; we threw them in voluntarily.

As things turned out, the main (and considerable) pay-off of this line of work came from anti-flea and anti-cockroach

formulations.

Djerassi's later life revolves around Stanford, the university and nearby research-and-development community, and an enthusiasm for sniffing out research opportunities in business. Most, but not all, were to do with pest control, and not all were successful (the account of a doomed foray into film making, from earlier times, is a hoot).

It's all urbanely done, by and large, with the single note of personal bitterness being reserved for his mother, of all



From the cover of *The Pill, Pygmy Chimps, and Degas' Horse*.

Carl Djerassi — from clean-cut chemist to bearded bard.

people. There is too a nice line in self-deprecation woven through the book. What isn't clear is what Djerassi's subordinates have made of him. He says he hung up his lab coat in 1952; is he a tyrant of the research group, a fixer or a mentor? Also not clear is why he didn't take the step into the ranks of the truly Good and Great (which is not to say he hasn't had his honours and influence in policy making and such). He himself points out that in American academic circles he has remained mostly an outsider. Too much the maverick maybe.

Perhaps I inhabit overly parochial social circles, but I find that a common response to mention of Djerassi's name is 'Carl who?'. This remarkable autobiography of a remarkable man should put an end to that sad state of affairs. □

Tim Lincoln is News and Views editor of Nature.

Market values

John L. Casti

Economics: Mathematical Politics or Science of Diminishing Returns? By Alexander Rosenberg. *University of Chicago Press: 1992. Pp. 266. \$37.50, £25.95.*

It has often been said that economics is either the hardest of the soft sciences or the softest of the hard sciences. In this provocative and important work, Alexander Rosenberg, professor of philosophy at the University of California, Riverside, claims that it is neither, for the simple reason that by whatever criteria you use to label an area 'scientific', economics fails to measure up.

How is one to judge whether economics, or any other field of study for that matter, is in any way scientific in its goals and methods? According to Rosenberg, one litmus test for the scientific is that a genuine science improves the predictive powers of its theories over time. Economics has failed dramatically by this standard, being limited to generic predictions of the sort exemplified by the statement made in 1989 by the investment house Drexel, Burnham, Lambert: "There are several alternative outcomes which are possible for the US economy in 1990". Rosenberg attributes a large part of the failure of economics to improve its predictive powers to the fact that the explanatory variables of folk psychology — beliefs, desires and preferences — that economists build their theories on "cannot be linked up with the rest of science. They cannot even be realized, exemplified, or instantiated by the brain in theoretically tractable ways." So if economics is not a natural science, what is it?

Perhaps it is a biological science. Pursuing this possibility, Rosenberg examines the biologically motivated theory of consumer behaviour advocated by Gary Becker, a theory that contributed greatly to Becker's winning this year's Nobel prize in economics. This theory draws parallels between an individual's economic needs and evolutionary biology, arguing that by focusing on populations rather than individuals, it is indeed possible to make meaningful predictions about aggregates such as market demand and supply. Sad to say, Rosenberg concludes that treating the theory of rational choice as a theory about biological systems holds out little hope of providing a useful line of development for economic theory.

As a final possibility, Rosenberg asks if economics is not an empirical science, then perhaps we should regard it as a branch of applied mathematics. Indeed,

both the defenders and detractors of economic theory have found its intellectual core in its mathematical expression. The central aspect of the 'mathematization' of economics is, Rosenberg claims, the identification of the intentionality of economic agents with the equilibria of, for example, a utility function.

Euclidean geometry was once styled the 'science of space', but calling it a science did not make it one. Drawing parallels between the development of economics as a mathematical activity and the development of geometry, the author notes that while we can certainly call economics the science of the distribution of scarce resources, this appellation again does not make it a science. Economists, like geometers, have blithely carried on their business of proving theorems without giving much thought to the issue of the cognitive status of economic theory. And the assumptions underlying the theorems of both euclidean geometry and neoclassical mathematical economics seem to rest on the same shaky ground. In this sense, the parallel postulate of Euclid and the assumption of economic rationality are of a kind. Rosenberg finally concludes that thinking of economics as mathematics enables us to explain a wide variety of the failings of economic theory to match up to economic facts.

The bottom line of Rosenberg's thesis

is that the best way to think of economics is not as a science, but rather as a combination of political philosophy and applied mathematics (with emphasis on the mathematics). Given the central role that economics seems to play in arriving at public policy decisions, this should be a sobering conclusion for all of us. This well argued and researched book should therefore be read by politicians, government mandarins and, more generally, the concerned lay public. But unfortunately its resolutely academic tone and rather dry literary style are unlikely to attract the attention of those for whom its message would do the most good. Perhaps some day someone somewhere will translate Rosenberg's arguments into plain everyday English and publish them in a form accessible to everyone. In the meantime, we can hope that intellectual osmosis will work its magic and Rosenberg's important message that economics is not and never can be a science will make its way from the rarefied halls of academia to the smoke-filled backrooms of Washington, Wall Street and Whitehall. □

John L. Casti is at the Institute for Econometrics, Operations Research and System Theory, Technical University of Vienna, Vienna, Austria, and the Santa Fe Institute, Santa Fe, New Mexico 87501, USA.

The sound of music

Michael Neve

The Volcano Lover: A Romance. By Susan Sontag. *Farrar, Straus and Giroux/Jonathan Cape: Pp.419. \$22, £14.99.*

THIS story centres on a famous *ménage à trois* of the late Enlightenment, the relationship between the British ambassador to Naples, Sir William Hamilton, a Fellow of the Royal Society, admirer of volcanoes and a dealer in paintings and Greek vases; Emma Lyon, the one time prostitute who became his mistress and his second wife; and Horatio Nelson, the naval hero and victim of Trafalgar. A great deal of Sontag's interest centres on the varieties of ways in which this trio and other characters observe the world and are observed in their turn. They observe Vesuvius, which constantly threatens eruptions that can both terrify and fascinate. Hamilton, more than 30 years older than Emma, and a man who had started collecting art objects in early life through an unhappy first marriage and an incapacity to get involved, is seen as a connoisseur of ruins. He is not looking, but looking away. This also allows him

to look at Vesuvius, but from a distance. When he meets his future wife Emma, he is physically excited by her but also has to add her to his art collection, putting her in different classical postures, such as that of Dido or Ariadne. Into the emptiness that sustains Hamilton's need to make living bodies into statues arrives the figure of Nelson. As a result, seven years after marrying Hamilton in 1791, Emma turns the connoisseur and diplomat into a kindly-regarded cuckold. The relationship between Emma and Nelson in its turn migrates into a form of gothic romance before Nelson assists in the destruction of Neapolitan republicanism and eventually dies heroically at Trafalgar. A portrait of modern Europe has been painted by looking at the lives of this trio.

Sontag's excursion into the world of the historical novel has aroused considerable interest, since she tells a story ostensibly set in Naples in the eighteenth century while managing to introduce comments on the modern world that are clearly her own. Part of the interest, therefore, lies in trying to work out what she is up to. What is the status of the

reflections on psychology, or the current state of New York streets, set in the historical context of the Bay of Naples in 1772? A great deal of one's sense of the success or failure of this book depends on one's answer to this question. If one feels preached at in the reflective passages, the book begins to lose its hold, especially if the lectures on, say, the nature of credibility or pleasure, or the effects of distance on one's appreciation of objects, seem facile. On the other hand, these interruptions of the modern voice often work wittily and engagingly, helping the book, in the best sense, to become a collector's item in itself, perhaps a picture from the past. Or, most ambitiously, an echo of that other late-eighteenth-century glimpse of the Bay of Naples, Mozart's *Così fan tutte*. Indeed, the theme of amorality, cynicism and fickleness is not one that this allegedly elitist and obscure writer makes difficult to grasp. In musical terms, one might even describe the subject of the book as a question of vocal emphasis: the volcano lover or the volcano lover?

The musical allusions are completed at the end of the novel with four monologues by women whose lives had interlocked with Hamilton's. The last of these is the Neapolitan revolutionary Eleonora de Fonseca Pimental, Portuguese by birth and an eventual convert to republicanism. She loses her life in the cause of anti-elitist politics but not before the book gives her its last words: "... I cannot forgive those who did not care about more than their own glory or well-being. They thought they were civilized. They were despicable. Damn them all."

Many readers will find the varieties of tone and of aesthetic judgement in this novel annoying, and no doubt further proof of the author's personal need to show off and be obscure. But few authors are as accustomed to such accusations as Sontag. By bravely carrying on regardless, and by drawing on a remarkable geological, sexual and political story, she has answered back. She describes a necessary tragedy with an anti-quarian's touch of gallows humour: throughout, we know that the volcano can stay dormant, allowing us to admire it, or explode, when we must run or be buried. The operatic intelligence that keeps this novel alive allowed this reader at least to hear the unmistakable sound of Maria Callas between the lines: she is never mentioned, but part of Sontag's purpose in introducing scenes from Puccini's *Tosca* is to hint at her reader's knowledge and let music do the rest. □

Michael Neve is at the Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BE, UK.

New in paperback

- *The Global Condition: Conquerors, Catastrophes and Community* by W. H. McNeill. Incorporates *The Great Frontier* (1983) and *The Human Frontier* (1979). Princeton University Press, \$12.95.
- *Handedness and Developmental Disorder* by D. V. M. Bishop. Lawrence Erlbaum Associates, £9.95.
- *The Science of Matter* edited by M. Crosland. An anthology of writings on physics and chemistry from the seventeenth to the twentieth centuries. Gordon and Breach, £18, \$34.
- *The Anthropology of Numbers* by T. Crump. Cambridge University Press, £11.95, \$15.95.

- *Brain Organization and Memory: Cells, Systems, and Circuits* edited by J. L. McGaugh, N. M. Weinberger and G. Lynch. Oxford University Press, £25.
- *Likelihood* by A. W. F. Edwards (expanded edition). A landmark volume on scientific inference, first published in 1972. Johns Hopkins University Press, \$15, £11.
- *Paul Broca: Explorer of the Brain* by Francis Schiller. A biography of the French surgeon-anthropologist, first published in 1979. Oxford University Press, £15, \$18.95.
- *Superstrings: A Theory of Everything?* edited by P. C. W. Davies and J. Brown. Canto (Cambridge University Press), \$9.95, £5.95.

Genetic drift

Christopher Wills

The Search for the Gene. By Bruce Wallace. Cornell University Press: 1992. Pp. 224. \$34.95 (hbk), \$12.95 (pbk).

I AM fond of saying that inside every fat book there is a thin one screaming to get out. This book is an exception — it is a thin book, but, defying logic, there is a fat one inside screaming to get out.

Bruce Wallace provides us with a highly personal, highly selective and — dare I say it — highly quirky survey of the search for the gene. He examines some of the more obscure and dusty byways of the history of genetics at great length, often with considerable insight and humour. But the book leaves out so much that the reader looks in vain for its shadow companion, the fat book that might have been. Bits of the fat book actually appear here and there, in the form of boxes that deal in more depth with various topics and together are almost as long as the main text.

The author begins with a quick look at the world of agriculture and of plant and animal breeding before Mendel, and quotes Haldane's comment that there was probably very little conscious selection going on during the appearance of domesticated crops — that selection for the most fertile seeds would have happened largely automatically. Then, two pages later, he points out that farmers must have been very observant to find the first few plants of the various strains of wheat that produce hard and soft flours. The reader expects him to pick a fight with Haldane, which would be easy enough, but instead he moves on.

A few pages further, tyro readers will be bewildered by an extended discussion of Boveri's important work on dispermy in sea-urchin eggs, which showed as

early as 1902 that each chromosome is essential for development. It would have been a good deal clearer if the difference between three-poled and four-poled dispermic eggs had been explained at the outset, but without this foundation I expect most readers will soon be lost.

Much later, Wallace takes the reader through an interesting history of how the genetic code was worked out. He quite properly emphasizes something that is usually glossed over: that by 1963 most of the code had been worked out by clever use of artificial oligomers of RNA. But he does not mention the beautiful demonstration by Crick, Barnett, Brenner and Watts-Tobin, done at the start of the search for the code, that the code must be read in threes or, less probably, in multiples of three. Nor does he discuss the work by Nirenberg, Leder and others that rapidly and elegantly clinched the identity of all the codons and essentially established the code almost overnight. The story is disconcertingly incomplete, emphasizing things that are usually left out and leaving out things that are usually emphasized.

The audience for the book is not clear. It might be read by undergraduates with a little spare time (though I haven't met any of those for quite some time) who are taking a genetics course and want to explore some of the background to what is certainly one of the most exciting scientific stories of our century. Readers who do have that background will certainly stumble on things they did not know, and will enjoy Wallace's refreshing approach to things we take for granted. The book is really an old-fashioned ramble rather than a tightly organized history, and should have been presented as such. □

Christopher Wills is in the Department of Biology, University of California, San Diego, La Jolla, California 92093-0116, USA.

A synaptic model of memory: long-term potentiation in the hippocampus

T. V. P. Bliss & G. L. Collingridge

Long-term potentiation of synaptic transmission in the hippocampus is the primary experimental model for investigating the synaptic basis of learning and memory in vertebrates. The best understood form of long-term potentiation is induced by the activation of the *N*-methyl-D-aspartate receptor complex. This subtype of glutamate receptor endows long-term potentiation with Hebbian characteristics, and allows electrical events at the postsynaptic membrane to be transduced into chemical signals which, in turn, are thought to activate both pre- and postsynaptic mechanisms to generate a persistent increase in synaptic strength.

THE assumption that information is stored in the brain as changes in synaptic efficiency emerged about a century ago following the demonstration by Cajal that networks of neurons are not in cytoplasmic continuity but communicate with each other at the specialized junctions which Sherrington called synapses. External events are represented in the brain as spatio-temporal patterns of neural activity, and it is these patterns of activity which must themselves be the agents of synaptic change. The location of storage, the engram of learning and memory, must therefore be found among those synapses which support activity-dependent changes in synaptic efficiency. These ideas were refined in the late 1940s by Hebb¹ and Konorski², who proposed a coincidence-detection rule in which the synapse linking two cells is strengthened if the cells are active at the same time. The first such synapses to be identified in the mammalian brain were the excitatory connections made by perforant path fibres onto granule cells of the hippocampus, a cortical structure required for the formation of conscious memories in man. Brief trains of high-frequency stimulation to monosynaptic excitatory pathways in the hippocampus cause an abrupt and sustained increase in the efficiency of synaptic transmission. This effect, first described in detail in 1973^{3,4}, is called long-term potentiation (LTP). LTP has since been found in all excitatory pathways in the hippocampus, as well as in several other regions in the brain, and there is growing evidence that it underlies at least certain forms of memory^{5,6}. In the past 10 years, LTP in the hippocampus has become the dominant model of activity-dependent synaptic plasticity in the mammalian brain, and much progress has been made in elucidating the mechanisms underlying its induction and expression.

Properties of hippocampal LTP

Activity-dependent synaptic potentiation occurs within milliseconds and can persist for many hours in the anaesthetized animal or in the *in vitro* hippocampal slice preparation, and for days when induced in the freely moving animal. This time span incorporates a number of mechanistically distinct temporal components, which include post-tetanic potentiation (PTP), short-term potentiation (STP) and LTP. Activity-dependent potentiation can also be classified on the basis of whether or not its induction is blocked by antagonists of the *N*-methyl-D-aspartate (NMDA) subtype of glutamate receptor (Box 1). In this article, by LTP we mean synaptic potentiation, which is both NMDA receptor-dependent and lasts for more than an hour.

LTP is expressed as a persistent increase in the size of the synaptic component of the evoked response, recorded from individual cells or from populations of neurons. It can be induced in a number of ways, most conveniently by delivering a tetanus (typically a train of 50–100 stimuli at 100 Hz or more) to the pathway of interest (Fig. 1). LTP can also be induced by more modest stimulus parameters, providing the patterns of

stimulation fall within certain critical ranges. (Two particularly efficient recipes are termed 'theta-burst stimulation'⁷ (for example, several bursts of 4 shocks at 100 Hz delivered at an interburst interval of 200 ms) and 'primed-burst stimulation'⁸ (for example, a single priming stimulus followed at 200 ms by a single burst of 4 shocks at 100 Hz). The significance of these protocols is that synchronized firing patterns at similar frequencies occur in the hippocampus during learning⁹.)

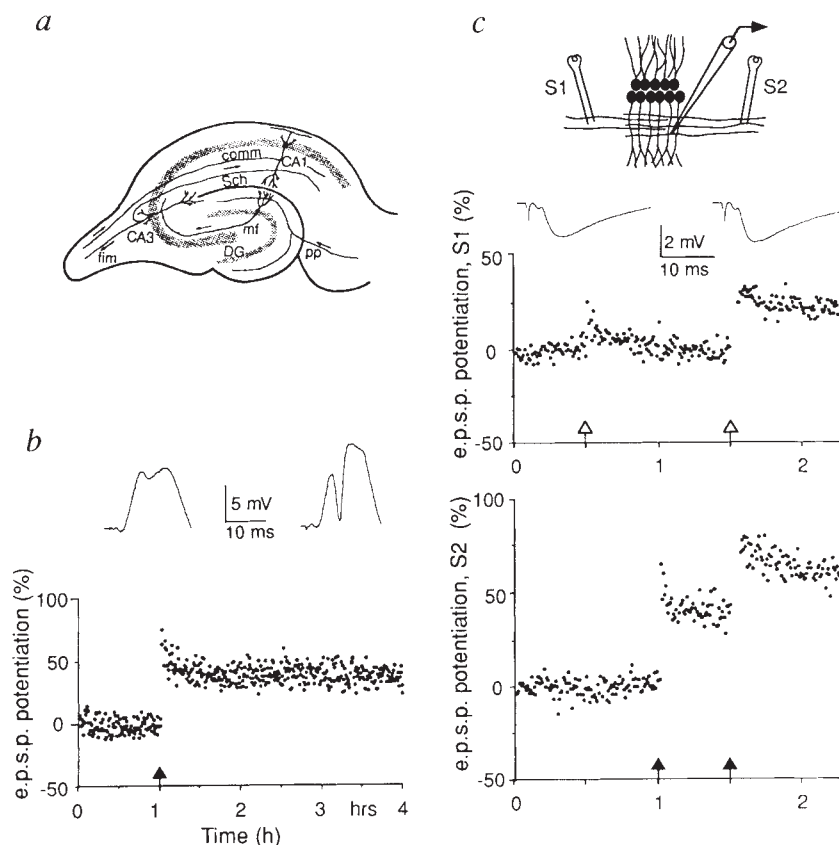
LTP is characterized by three basic properties: cooperativity, associativity and input-specificity. Cooperativity describes the existence of an intensity threshold for induction; 'weak' tetani, activating relatively few afferent fibres, do not trigger LTP¹⁰. The threshold for inducing LTP is a complex function of the intensity and pattern of tetanic stimulation; between 'weak' trains which produce only PTP and 'strong' trains which induce LTP, lies an intermediate range of activation which engages STP^{11,12}. LTP is associative in the sense that a 'weak' input can be potentiated if it is active at the same time as a strong tetanus to a separate but convergent input^{10,13}. Finally, LTP is input-specific, because other inputs that are not active at the time of the tetanus do not share in the potentiation induced in the tetanized pathway^{14,15}. Associativity provides a cellular analogue of classical conditioning, and is an implicit property of the Hebb synapse, the computing element that lies at the heart of the current interest in neural computation. The three properties can be explained on the assumption that a synapse will be potentiated if, and only if, it is active at a time when the region of dendrite on which it terminates is sufficiently depolarized. Validation of this induction rule was provided in 1986 by experiments showing that low-frequency (1 Hz), low-intensity stimuli could produce robust LTP if repeatedly paired with depolarizing pulses delivered through an intracellular recording electrode^{16–18}. In the limit, LTP can be produced in this way between pairs of synaptically coupled neurons¹⁹. Conversely, the induction of LTP can be blocked by limiting the depolarization of the cell during a tetanus^{16,20}.

What is now needed to complete a mechanistic description of the induction requirements for associative LTP is a molecular coincidence detector, able to respond to the conjunction of activity in afferent fibres and adequate depolarization in target dendrites. Compelling evidence that the NMDA receptor performs this function is reviewed in the next section.

The induction of LTP

The role of amino-acid receptors in the induction of LTP. The involvement of several amino-acid receptor subtypes in the induction of LTP has been determined largely by the use of antagonists and is described in Box 2. The key role of the NMDA receptor channel complex relies on several of its special properties, in particular the voltage-dependent block of its channel by Mg²⁺ (ref. 21). It is this that allows the NMDA receptor to

FIG. 1 Basic properties of LTP: cooperativity, input-specificity and associativity. *a*, Simplified diagram of a transverse section through the hippocampus of the rat, showing the principal neuronal fields (granule cells of the dentate gyrus (DG) and the pyramidal cells of areas CA3 and CA1), and the main excitatory afferent projections (the perforant path (pp) from entorhinal cortex to granule cells, the mossy fibre projection (mf) from granule cells to CA3 cells, and the Schaffer collateral (Sch)-commissural (comm) system which connects ipsilateral and contralateral CA3 cells to CA1 cells). Interneurons, which are found in all hippocampal subfields and which form powerful inhibitory connections with principal cells through feed-forward and feedback loops, have been omitted. *b*, An example of LTP in the perforant pathway recorded *in vivo*. The graph plots the slope of the rising phase of the evoked response (population e.p.s.p.), recorded from the cell body region in response to constant test stimuli, for 1 h before and 3 h following a tetanus (250 Hz, 200 ms), delivered at the time indicated by the arrow. Representative traces before and after the induction of LTP are illustrated above the graph. Note the increase in slope of the population e.p.s.p. and the increase in size of the superimposed population spike (downward deflection). *c*, Demonstration of the properties of cooperativity, input specificity and associativity. The diagram at the top shows the experimental arrangement in area CA1 of the hippocampal slice preparation. Two independent sets of afferent fibres converging on a common population of cells are activated by stimulating electrodes (S1 and S2) placed either side of the extracellular recording electrode. The stimulus intensities are adjusted so that S1 activates fewer fibres than S2. The slope of the population e.p.s.p.s, in response to stimuli delivered alternately to S1 and S2 at 15-s intervals, are plotted as a function of time. Arrows denote episodes of tetanic stimulation to S1 (the 'weak' pathway, open arrows) or S2 (the 'strong' pathway, solid arrows). The tetanus to S1 produced a rapidly decaying phase of PTP, lasting 2–3 min, with a small tail of STP, but no stable increase in synaptic transmission; the intensity of the tetanus was below the cooperativity threshold for LTP. The stronger tetanus to S2 (first



filled arrow) produced PTP and robust LTP, but there was no transfer of the effect to the first input (test shocks to S1 were out of phase with the high-frequency bursts to S2), demonstrating the input-specificity of LTP. Finally, tetani to S1 and S2 were delivered together. The coincident activation of a weak, subthreshold input with a strong input induced associative LTP in the weak input. The traces above the graph illustrate field e.p.s.ps, evoked by test shocks in S1 and recorded in the synaptic layer, before and after the induction of associative LTP.

filled arrow) produced PTP and robust LTP, but there was no transfer of the effect to the first input (test shocks to S1 were out of phase with the high-frequency bursts to S2), demonstrating the input-specificity of LTP. Finally, tetani to S1 and S2 were delivered together. The coincident activation of a weak, subthreshold input with a strong input induced associative LTP in the weak input. The traces above the graph illustrate field e.p.s.ps, evoked by test shocks in S1 and recorded in the synaptic layer, before and after the induction of associative LTP.

behave as a molecular coincidence detector. For the NMDA channel to open, and thus to trigger the induction of LTP, it is necessary for two events to occur simultaneously: the membrane must be sufficiently depolarized to expel Mg^{2+} from NMDA channels at the same time that L-glutamate has, by binding to NMDA receptors, promoted their opening. The slow time course and voltage-dependence of the NMDA receptor-mediated conductance makes it particularly susceptible to the hyperpolarizing influence of synaptic inhibition²²; this susceptibility, together with the frequency-dependent depression of inhibition itself, largely accounts for the frequency-dependence of the induction of LTP²³.

The properties of cooperativity, associativity and input-specificity can now easily be explained. The cooperativity threshold follows from the need for depolarization to reduce the level of the Mg^{2+} block of the NMDA channel. 'Weak' stimuli, activating only a few fibres, fail to induce LTP not because insufficient L-glutamate is released to activate NMDA receptors, but because the level of depolarization provided by the weak input does not produce an adequate reduction of the Mg^{2+} block. When many fibres are activated in synchrony by a 'strong' stimulus, depolarization spreads between neighbouring synapses to enhance the unblocking of NMDA channels. Associativity has a similar explanation except that the required depolarization is provided by a different set of afferent fibres; in theory, these

'helper' inputs could use any neurotransmitter that promotes depolarization, and, experimentally, depolarization is often provided by injecting current into the cell. Input-specificity is explained by the need for the presynaptic terminal to provide a sufficient concentration of L-glutamate to activate adequate numbers of NMDA receptors. (It follows that there can be little activation of NMDA receptors by ambient or spontaneously released L-glutamate, otherwise LTP would be induced by depolarization alone.)

Because the induction of LTP by tetanic stimulation is prevented by a variety of NMDA antagonists, including those which act at the receptor (such as 2-amino-5-phosphonopentanoate (AP5)²⁴), in the channel (for example, MK-801 (ref. 25)) and at the allosteric glycine site (for example, 7-chlorokynurenic acid²⁶), it is clear that activation of these receptors is an essential trigger for the process. But, the application of NMDA itself is not usually sufficient to induce LTP, though it readily induces STP^{24,27}. A possible reason for this relates to the paradoxical finding that a level of activation of the NMDA receptor system, which is itself inadequate for producing LTP, can result in a subsequent impairment in the ability to generate LTP^{28–30}. Thus, with the application of NMDA there may be two opposing processes at work, one promoting and the other suppressing the induction of LTP. Alternatively, factors in addition to NMDA receptor activation, which could be either pre- or post-synaptic,

may be required to facilitate or allow the induction of LTP. In this respect, there has been interest in the possible role of metabotropic glutamate receptors (mGluRs), prompted by the observation that the mGluR antagonists 2-amino-4-phosphonobutanoate (AP4) and 2-amino-3-phosphonopropionate (AP3) reduce the duration of LTP^{31,32}. These compounds are, however, very weak mGluR antagonists of poorly defined specificity, and the observation will need to be confirmed as soon as more potent and selective mGluR antagonists are developed. A second indication that these receptors might be involved in the induction of LTP has come from the finding that aminocyclopentane dicarboxylate (ACPD), the 1S,3R-enantiomer of which is a specific agonist for mGluRs, can augment tetanus-induced potentiation³³. In addition, ACPD enables sub-threshold³⁴, or low-frequency stimuli (in conjunction with the application of NMDA)³⁵, to induce LTP. It does this in at least two ways. First, ACPD augments responses of hippocampal neurons to NMDA³⁴. Second, it can elicit an NMDA receptor-independent potentiation of slow onset which adds to STP to produce a potentiation that closely resembles tetanus-induced LTP³⁶.

The role of Ca^{2+} in the induction of LTP. In an important early study, it was found that the induction of LTP could be blocked by the intracellular injection of the Ca^{2+} chelator EGTA³⁷. This result implicated the postsynaptic cell, and in particular Ca^{2+} signalling in the induction process. Because NMDA channels are permeable to Ca^{2+} (refs 21, 38, 39) it is widely assumed, but not proven, that permeation through these channels during tetanic stimulation provides the Ca^{2+} signal necessary for the induction of LTP. Because NMDA receptors are assumed to be located on dendritic spines, it is believed that spines may act to localize the Ca^{2+} signal. Spines can restrict the diffusion of Ca^{2+} (ref. 40); however, whether they do so in LTP is not known.

Using Ca^{2+} -imaging techniques it has been shown that tetanic stimulation elevates Ca^{2+} within dendrites and spines^{41,42}. Part

of this signal depends on the synaptic activation of NMDA receptors and reflects, at least in part, Ca^{2+} entry through NMDA channels and voltage-gated Ca^{2+} channels. In one study⁴² the tetanically induced rise in Ca^{2+} persisted for several minutes, and it was proposed that sustained Ca^{2+} gradients might be important for memory processing. But it is unlikely that rises in Ca^{2+} of this duration are necessary for the induction of LTP in view of the demonstration that LTP can still be induced even if the duration of the post-tetanic rise in Ca^{2+} is restricted to less than 3 s, using a photo-activatable caged Ca^{2+} chelator⁴³. Complementary data have come from combining Ca^{2+} imaging with whole-cell recording⁴⁴. Although, for technical reasons, LTP could not be induced, this preparation allowed Ca^{2+} signals to be correlated directly with the synaptic response. Strong tetanic stimulation, which evoked large NMDA receptor-mediated synaptic currents, produced Ca^{2+} transients lasting only a few seconds. This combination of techniques has also enabled the Ca^{2+} signal that permeates NMDA channels on dendritic spines to be detected (Fig. 2).

There are indications from Ca^{2+} imaging experiments that the Ca^{2+} which permeates NMDA channels is augmented by Ca^{2+} release from intracellular stores (see Box 3). The Ca^{2+} transient associated with the synaptic activation of NMDA receptors is substantially reduced in the presence of ryanodine or thapsigargin⁴⁴, drugs which inhibit Ca^{2+} -induced Ca^{2+} release and deplete intracellular Ca^{2+} stores, respectively. That this Ca^{2+} might be important for the induction of LTP is suggested by the observations that dantrolene, which acts at the ryanodine receptor, and thapsigargin can both inhibit the induction of LTP^{36,45,46}. It is likely that inositol 1,4,5-trisphosphate (InsP_3) generated as the result of the activation of mGluRs, as well as the Ca^{2+} which permeates through NMDA channels, is involved in releasing Ca^{2+} from intracellular stores. Moreover, activation of mGluRs can induce LTP by a thapsigargin-sensitive mechanism, even if NMDA receptors are blocked³⁶. This suggests that

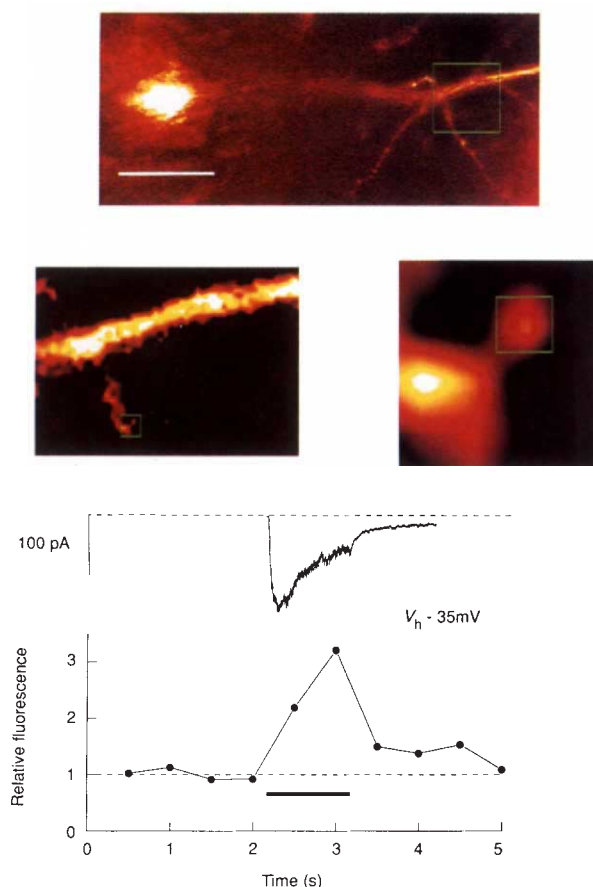


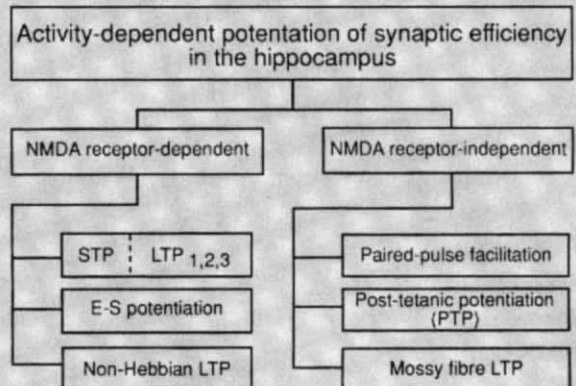
FIG. 2 Ca^{2+} permeates NMDA channels to produce a transient signal in spines in response to tetanic stimulation. Confocal images of a CA1 pyramidal neuron in a hippocampal slice. The upper image shows part of the soma and dendrites as they emerge into the plane of the optical section. The boxed region is enlarged to show a dendritic branch and spine-like structure. This is further enlarged to show the spine in more detail and the boxed region from which the fluorescence measurements were obtained. (The box is $\sim 1 \mu\text{m}^2$.) The graph plots the relative fluorescence, emitted by the indicator fluo-3, as a function of time. The tetanus (100 Hz, 1s), delivered for the duration of the bar, resulted in a transient increase in fluorescence. The upper trace shows the synaptic current induced by the tetanus, recorded through a patch-pipette. The cell was internally dialysed and clamped at -35 mV to eliminate all voltage-gated Ca^{2+} channel activity and the slice was treated with thapsigargin to deplete intracellular Ca^{2+} stores. Under these conditions the fluorescence changes are caused by Ca^{2+} permeating through NMDA channels.

BOX 1 Classification of activity-dependent increases in synaptic efficiency in the hippocampus

SYNAPTIC potentiation can be divided into two principal categories on the basis of whether or not its induction is blocked by antagonists of the NMDA subtype of glutamate receptor. Several categories of NMDA receptor-dependent plasticity have been identified. A distinction can be made between short-term potentiation (STP), which decays within 1 h, and long-term potentiation (LTP), which is sustained for much longer periods. STP can be distinguished from LTP by the use of protein kinase inhibitors, in the presence of which potentiation usually persists for only 30–60 min^{81–85}. Potentiation of a similar duration can be produced by decreasing the number of stimuli in the tetanus or by other manoeuvres which reduce the level of NMDA receptor activation¹². Although it is convenient to make the distinction, the relationship between STP and LTP has not been clearly defined. LTP can be tentatively subdivided into several mechanistically distinct components: LTP1, with a duration of less than 3–6 h which is blocked by kinase inhibitors but not by protein synthesis inhibitors; LTP2, a component which is blocked by translational inhibitors but which appears to be independent of gene expression; and LTP3, with a time constant of several days, which is only obtained if the animal is anaesthetised at the time of induction¹⁵⁹ and which may require gene expression (see text).

Another form of NMDA receptor-dependent plasticity is E–S potentiation. This takes its name from the shift to the left of the curve relating the slope of the population e.p.s.p. (E) to the amplitude of the population spike (S) which is commonly observed following a tetanus¹⁶⁰. It appears not to be input-specific¹⁶¹ but may provide a generalized boost to impulse traffic. A further type of presumed NMDA receptor-dependent LTP has been described in which potentiation occurs not only at those synapses where there is coincident pre- and postsynaptic activity, but extends to synapses made by concurrently active terminals onto neighbouring cells, whether or not these are active¹⁶². This is interesting both because it suggests that non-Hebbian forms of potentiation occur in the hippocampus, and because it provides implicit evidence for the existence of a diffusible extracellular messenger (see text).

NMDA receptor-independent processes include paired-pulse facilitation and post-tetanic potentiation (PTP), which are general features of excitatory synaptic transmission. With the stimulus parameters usually employed to produce LTP, the duration of PTP is at most a few minutes. Both paired-pulse facilitation and PTP are additive with LTP, and can be produced repeatedly even when LTP has reached asymptotic levels. It follows that LTP cannot achieve the maximum strength of which a synapse is capable:



the potential for a further short-term increase is always held in reserve. Mossy fibres terminate in the stratum lucidum of area CA3, a subfield devoid of NMDA receptors. Consistent with this observation, LTP in mossy fibres is not blocked by the NMDA antagonist AP5 (ref. 163); moreover, it appears to be nonassociative. The projection is technically difficult to study, and the locus and cellular mechanisms of mossy fibre LTP remain controversial¹⁶⁴. Finally, an input-specific AP5-resistant component of LTP has been described in area CA1¹⁶⁵. The effect is small, develops gradually, is blocked by Ca^{2+} channel antagonists, and requires stronger tetanic stimulation for its induction than is needed for NMDA receptor-dependent LTP.

Long-lasting potentiation can also be induced by transient exposure of hippocampal synapses to a variety of chemical agents, including Ca^{2+} ¹⁶⁶, arachidonic acid⁶⁶, the metabotropic glutamate receptor (mGluR) agonist aminocyclopentane-1S,3R-dicarboxylate (1S,3R-ACPD)³⁶, the K^+ channel blocker, tetraethylammonium (TEA)¹⁶⁷ and the G-protein activator NaF/AlCl_3 ¹⁶⁸. Chemically-induced potentiation usually occludes with tetanically-induced LTP (that is, saturation of one prevents induction of the other), suggesting a convergence of mechanisms; in general, chemically induced LTP is not blocked by NMDA antagonists, presumably because the components of the LTP cascade activated by the various agents lie downstream from the NMDA receptor.

release of Ca^{2+} from intracellular stores can substitute for the NMDA receptor-mediated Ca^{2+} signal. Other routes by which Ca^{2+} could enter the cell to contribute to the induction of LTP include voltage-dependent Ca^{2+} channels and Ca^{2+} -permeable AMPA channels (that is, those lacking the GluR-2 subunit⁴⁷). At present, though, there is little evidence that either of these pathways plays a significant role in LTP.

Although clearly a necessary factor, it is unclear whether a rise in postsynaptic Ca^{2+} provides a sufficient trigger for the induction of LTP. Elevation of intracellular Ca^{2+} by the photolysis of caged Ca^{2+} induces a form of synaptic potentiation⁴⁸, but the relationship between this effect and LTP has not been determined (for example, occlusion experiments have not been done). Elevation of intracellular Ca^{2+} , either by evoking Ca^{2+} currents⁴⁹ or by slowly depleting intracellular Ca^{2+} stores⁴⁶, does not induce LTP. This could be due either to the failure of these methods to elevate Ca^{2+} in the appropriate manner (presumably what is needed is a large transient within spines) or to the need for additional pre- and/or postsynaptic signals.

In summary, the available evidence suggests that under normal conditions Ca^{2+} permeates NMDA channels to provide a transient signal which is necessary for the induction of LTP. It is probable that this signal is restricted to the vicinity of activated spines and is amplified by release from intracellular stores.

Expression of LTP

A major challenge is to identify the loci and nature of the alterations responsible for the expression of the potentiated state.

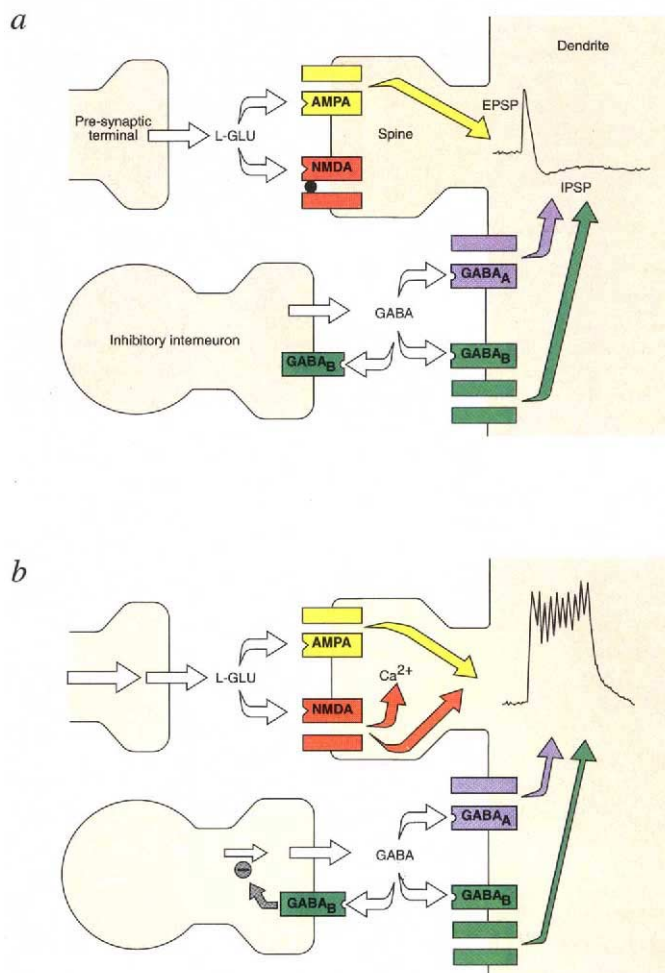
The locus of expression. Broadly speaking, the increase in the postsynaptic response generated at potentiated synapses could be due to (1) presynaptic modifications which result in an increase in the amount of L-glutamate released per impulse, (2) postsynaptic modifications, such as an increase in the number of receptors or a change in their functional characteristics, (3) an extrasynaptic change, such as a reduction in uptake of L-glutamate by glial cells leading to increased neurotransmitter availability at the receptors, or (4) morphological modifications. In reality, a combination of these changes, with different time courses, probably occurs.

Evidence for an increase in neurotransmitter release is derived from experiments that have measured the overflow of radiolabelled or endogenous L-glutamate from the hippocampus before and after the induction of LTP^{50–52}. Although not without difficulties of interpretation⁵³, these experiments establish a case for presynaptic changes lasting for at least several hours⁵⁴. Supporting evidence was obtained from experiments in which the ability of a depolarizing stimulus to release radiolabelled glutamate was shown to be elevated in potentiated hippocampal tissue⁵⁵.

Other studies have suggested purely postsynaptic modifications. The observation that paired-pulse facilitation is not altered after the induction of LTP has been interpreted as evidence for a postsynaptic modification in LTP⁵⁶ on the assumption that facilitation in the hippocampus is presynaptic, and that an interaction between facilitation and LTP would be expected if the expression of the latter were also presynaptically mediated. It is possible, however, to construct a model in which

BOX 2 The role of amino-acid receptors in the induction of LTP

a, Low frequency transmission: a single stimulus applied to the Schaffer collateral-commissural pathway evokes an e.p.s.p. which is mediated predominantly by the neurotransmitter (L-GLU) acting on ionotropic glutamate receptors of a non-NMDA type²⁴. This e.p.s.p. can be blocked by the quinoxalinedione antagonists, such as 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX)^{169,170}, and is usually referred to as AMPA receptor-mediated after the selective ligand for these receptors α -amino-3-hydroxy-5-methyl-4-isoxazolepropionate (AMPA). This receptor corresponds to the cloned family GluR1-4⁴⁷. When the Schaffer collateral-commissural pathway is stimulated it also activates GABAergic interneurons (through glutamatergic synapses similar to those on pyramidal neurons¹⁶⁹) and this leads to the biphasic i.p.s.p. which curtails the e.p.s.p. The initial part of the i.p.s.p. is caused solely by the activation of GABA_A receptors (which contain integral Cl⁻ channels) and this is supplemented and followed by the activation of GABA_B receptors (which are indirectly coupled to K⁺ channels). NMDA receptors contribute little to the synaptic response because of their relatively slow activation kinetics^{22,171,172}. By the time that significant numbers of NMDA channels are in an open state the neuron has been hyperpolarized by the i.p.s.p. and this greatly enhances the block of NMDA channels by Mg²⁺ (ref. 22). Even so, there will still be a finite contribution of the NMDA receptor system to low frequency synaptic transmission; however, this is not sufficient (under normal circumstances) to initiate changes in the efficiency of synaptic transmission. **b**, High-frequency transmission. The contribution of NMDA receptors to synaptic transmission alters radically in response to a high-frequency input^{24,173}. This is because the tetanus maintains the neuron in a more depolarized state, which in turn reduces the extent of the Mg²⁺-induced block of NMDA channels, while at the same time providing the L-glutamate which promotes their opening. Several factors may contribute to the sustained depolarization during a tetanus; these include summation of AMPA receptor-mediated e.p.s.p.s, depolarizing shifts in the Cl⁻ and K⁺ reversal potentials due to build up of intracellular Cl⁻ and extracellular K⁺. The primary mechanism (during primed or theta-burst LTP) is depression of GABA-mediated synaptic inhibition²³. This is an active process mediated by GABA_B autoreceptors. The effect takes more than 10 ms to develop and can last for up to a few seconds. As a result low frequency transmission is unaffected by this process; however, during high-frequency transmission there is considerably less GABA released per impulse which leads to a shift in the balance of excitation and inhibition. The reduction in inhibition allows greater expression of the NMDA receptor system which in turn contributes to the depolarization and thus further reduces the level of the Mg²⁺ block. The long duration of the synaptic conductance means that NMDA receptor-mediated e.p.s.p.s summate very effectively during high-frequency transmission.



facilitation and LTP are both presynaptic and yet involve additive, non-interacting mechanisms; this could be the case, for example, if the initial probability of release were very low. Claims that LTP is associated with a specific^{57,58} increase in the AMPA receptor-mediated component of the synaptic response, have formed the basis of an argument for a purely postsynaptic change, on the assumption that a presynaptic change would result in a similar increase in both AMPA and NMDA receptor-mediated components. In support of this argument, an increase in both components was seen during PTP, whereas the isolated NMDA receptor-mediated component failed to exhibit LTP. But the argument has been undermined by subsequent reports that NMDA receptor-mediated synaptic transmission exhibits pronounced LTP⁵⁹⁻⁶².

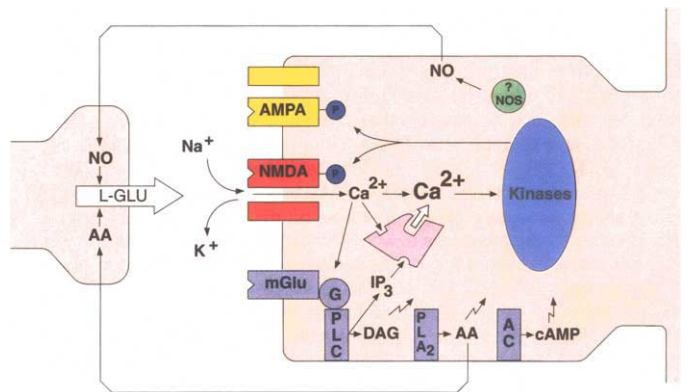
Another test for postsynaptic changes is to monitor the sensitivity of neurons to the application of agonists before and after the induction of LTP. Initial studies found no increase in the sensitivity to L-glutamate for up to 30 min post-tetanus^{63,64}. But in a more recent study, in which AMPA or quisqualate were used as agonists to avoid problems associated with the uptake and possible nonspecific actions of L-glutamate, a slow-onset increase in sensitivity was detected⁶⁵. The effect began within a few minutes but took an hour or more to reach a maximum. This time course parallels the slow-onset potentiation that can

be induced by the application of certain agents, such as arachidonic acid⁶⁶, and ACPD³⁶. It seems reasonable to assume that changes in the steady-state responses, as measured in the above experiments, reflect alterations that would also affect the response to synaptically released L-glutamate (for example, changes in the number, or conductance properties, of AMPA receptors). If this assumption is valid, then the results imply that the expression of STP is presynaptic whereas that of LTP is, at least in part, postsynaptic.

Despite the formidable interpretational problems of applying quantal analysis to central synapses, there has been a resurgence of interest in the use of this technique to analyse the locus of expression of LTP. Early studies in area CA1 indicated a presynaptic locus⁶⁷. Results of the more recent studies of fluctuations in the amplitude of synaptic responses have produced conclusions ranging from purely presynaptic⁶⁸, to predominantly presynaptic^{19,69,70}, to purely postsynaptic⁷¹ and, finally, to a mixture of purely presynaptic, purely postsynaptic and both pre- and postsynaptic^{72,73}. This variability may reflect differences in the initial release probability which, in turn, will be influenced by experimental parameters such as the extracellular Ca²⁺ concentration⁷³. Analysis of spontaneous miniature synaptic currents, associated with NMDA- or L-glutamate-induced STP, has provided evidence for an increase in quantal

BOX 3 Ways in which L-glutamate through its action on postsynaptic receptors may affect signal transduction processes involved in LTP

THE initial induction signal is a Ca^{2+} transient which permeates NMDA channels. This signal is then amplified by the release of Ca^{2+} from Ca^{2+} /InsP₃-sensitive intracellular stores. A parallel pathway which may be important for the induction of LTP is provided by mGluRs. These receptors can couple, through G-proteins, to the phosphoinositide-specific phospholipase C (PLC), phospholipase A₂ (PLA₂) and adenylate cyclase (AC)¹⁵⁵, to produce diacylglycerol (DAG), arachidonic acid (AA), and to regulate the levels of cAMP, respectively. Note that the initial NMDA receptor-mediated Ca^{2+} transient may be necessary for the activation of these mGluR cascades by L-glutamate¹⁴⁸. The amplified Ca^{2+} signal, in association with the other activators of protein kinases (zig-zag arrows), then leads to the phosphorylation of substrate proteins including, probably, AMPA and NMDA receptors. Other enzymes, such as nitric oxide synthase (NOS), if present, may also be activated by the Ca^{2+} transient. Biochemical changes in the presynaptic terminal may be initiated by the action of retrograde messengers, such as arachidonic acid (AA), nitric oxide (NO) and K^{+} , perhaps in conjunction with the action of L-glutamate on presynaptic mGluRs¹⁷⁴.



size in the hippocampal slice, implying a postsynaptic locus⁷⁴, and an increase in miniature frequency in cultured hippocampal neurons, implying a presynaptic locus⁷⁵. Evidently, the hoped-for resolution of the locus of expression of LTP by the application of quantal analysis has not yet been achieved. Note that if STP and the several temporal phases of LTP (see Box 1) are expressed at different loci, then changes in quantal parameters may alter progressively with time^{69,70}.

Signal transduction mechanisms. Several different Ca^{2+} -sensitive enzymes have been proposed to play a part in converting the probable induction signal, the entry of Ca^{2+} through the NMDA channel, into persistent modifications of synaptic strength. These include the protease calpain⁷⁶, phosphatases such as calcineurin⁷⁷, phospholipases and protein kinases. Most interest has focused on phosphorylation cascades and, in particular, the role of protein kinases. The first kinase to be implicated in LTP was the Ca^{2+} /phospholipid-dependent protein kinase (PKC)⁷⁸⁻⁸⁰. Inhibitors of the enzyme invariably block the induction of LTP; in most studies, STP is unaffected by PKC inhibitors⁸¹⁻⁸⁵, though with the use of high doses or the combined application of inhibitors STP may also be blocked⁸⁶. There is general agreement that PKC inhibitors will block LTP if they are applied after the tetanus, indicating that kinase activity outlasts the initial induction signal. But the duration of the time-window during which kinase inhibitors are effective and the manner in which the activation of kinases is maintained are both matters of debate. For example, it has been suggested that constitutively activated PKC is involved because H-7, which inhibits the activity of the catalytic subunit, but not sphingosine, which prevents the initial activation of PKC, can depotentiate synapses in a reversible manner even when applied up to 3 h after induction⁸³. But the selectivity of H-7 for potentiated pathways has been challenged⁸⁷, and other PKC inhibitors that act on the catalytic subunit, including K-252b (ref. 85), are not able to depotentiate fully established LTP. There is also disagreement as to whether the sustained kinase activity that might be necessary for LTP is located within the postsynaptic cell⁸⁶ or not⁸⁸. A recent view⁸⁹ is that a postsynaptic kinase is activated transiently (for less than a few minutes following the tetanus) and a presynaptic kinase is activated for longer periods (but for less than 1 hour). These kinases might be the γ and β isoforms of PKC, respectively. Intracellular injection of the catalytic subunit of PKC induces synaptic potentiation⁹⁰ as does the extracellular application of activators of PKC, such as certain phorbol esters⁹¹. But the enhanced response does not survive washout of phorbol ester, and occlusion experiments indicate that LTP and phorbol ester-induced potentiation use different mechanisms^{92,93}. Overall, it seems that activation of PKC is not

sufficient to induce LTP but is a necessary factor and may be specifically involved in the conversion of STP to LTP1 (that is in the consolidation or stabilization of LTP). The development of more selective PKC inhibitors and, in particular, subtype-specific inhibitors are needed to confirm and extend these ideas.

Several inhibitor studies have also indicated a role for calmodulin and the Ca^{2+} /calmodulin-dependent protein kinase CaMKII in LTP^{84,88,94-96}. Knockout of the gene encoding α CaMKII, an isoform which is heavily enriched in postsynaptic densities, severely impairs, though it does not always completely block, the ability of slices to exhibit LTP⁹⁷. The autophosphorylated form of this enzyme does not require Ca^{2+} and as a result becomes constitutively active. This has led to the proposal that CaMKII can act as a form of molecular memory, recording the occurrence of a previous Ca^{2+} transient⁹⁸. But contrary to the predictions of this model, NMDA does not alter the proportion of Ca^{2+} -independent CaMKII in organotypic hippocampal cultures⁹⁹. Less is known about the role of other kinases in LTP. The level of cAMP is elevated in an NMDA receptor-dependent manner in LTP and this may indicate an involvement of cAMP-dependent protein kinase (PKA)¹⁰⁰. It has been suggested, on the basis of inhibitor studies, that protein tyrosine kinases (PTKs) are involved in LTP¹⁰¹, and it may be relevant that NMDA receptor activation leads to tyrosine phosphorylation of MAP-2 kinase¹⁰².

In addition to post-translational modification of existing proteins there is evidence that protein synthesis is also necessary for LTP. The extent to which protein synthesis inhibitors prevent LTP is variable, depending on the inhibitor used. Probably the clearest picture has emerged from the use of anisomycin, which inhibits translation of proteins from mRNAs. If present at the time of the tetanus anisomycin reduces the duration of LTP to 3-6 hours¹⁰³⁻¹⁰⁵. A similar rate of decay is seen if LTP is induced in synapses that have been surgically isolated from the major site of protein synthesis in the cell body layer¹⁰⁵. In contrast, actinomycin, which prevents the transcription of mRNAs from DNA, has no effect on this anisomycin-sensitive phase¹⁰⁴. Taken together, these results suggest that proteins synthesized from pre-existing mRNA are required for the maintenance of LTP during the first few hours (corresponding to LTP2 in the classification shown in Box 1). The identity of proteins which are up- or downregulated during this period are not known, but several have been separated on two-dimensional gels¹⁰⁶. It is also intriguing that an increase in protease activity has been detected in perfusates from the dentate gyrus following potentiation¹⁰⁷, raising the possibility that cleavage of proteins with extracellular domains, such as neural cell adhesion molecules (NCAMs), may contribute to synaptic remodelling in LTP.

The experiment with actinomycin suggests that for the first 3 hours or so LTP does not depend on gene transcription. This does not preclude the possibility that genes are normally transcribed at or shortly after the time of induction but exert their effects at later times. An example of gene transcription induced by tetanic stimulation is the NMDA receptor-dependent increase in mRNA for the immediate early gene *zif/268* (refs 108–110). There is also a transient expression of *c-fos*, but only if the animal is unanaesthetised at the time of induction^{111,112}, suggesting that the *c-fos* protein is necessary for the generation of the most persistent form of LTP (LTP3; see Box 1). Changes in the abundance of mRNAs for a number of proteins have recently been identified in single CA1 cells 30 min to 3 hours after tetanization¹¹³. The reported changes in message for protein kinases (CaMKII is upregulated and the β -isoform of PKC is downregulated) suggests that protein kinases may play a role in the late stages of LTP, in addition to their presumptive action during the early phases.

Postsynaptic modifications. It is likely that the postsynaptic component of the expression of LTP involves alterations in the number and/or properties of the ion channels that mediate synaptic transmission. In view of the evidence that protein kinases are involved in LTP, the simplest scheme is that the kinases directly phosphorylate these ion channels. Consistent with this possibility, the gradual increase in AMPA sensitivity following the induction of LTP is prevented by K-252b, a potent kinase inhibitor⁸⁵. In addition, cloned AMPA receptors have several consensus sequences for phosphorylation by various kinases⁴⁷. Finally, the catalytic subunit of PKA can directly increase AMPA receptor function^{114,115}.

The finding that each of the AMPA receptor subunits can exist in two alternatively spliced variants, termed flip and flop, with different conductance properties, raises the possibility that LTP reflects a change in the relative expression of the flip and flop variants¹¹⁶. Alternatively, it could involve a change in the relative expression of the different subtypes of AMPA receptor, GluR 1–4 (ref. 47). A third possibility is regulation of RNA editing¹¹⁷.

The drug aniracetam, which potentiates responses to AMPA¹¹⁸ by preventing desensitization^{119,120}, has been used to explore how AMPA receptors may be modulated in LTP. The underlying idea is that if LTP and aniracetam share common mechanisms then their effects should interact. The weight of evidence suggests little interaction^{118,119,121,122} indicating that aniracetam and LTP do not regulate AMPA receptor function in the same manner.

So far, studies have concentrated on how the AMPA receptor-mediated component of synaptic transmission may be modified in LTP. But the NMDA receptor-mediated component also exhibits robust LTP^{59–62}. Alterations in this component could provide a means by which synapses increase their plasticity, as well as their efficiency. As with the AMPA receptor-mediated component, LTP of the NMDA receptor-mediated component of synaptic transmission could involve increases in L-glutamate release and/or postsynaptic modifications. A mechanism for the latter possibility is suggested by the observation that NMDA receptor function can be increased by the activation of PKC¹²³. This may involve phosphorylation of NMDA channels to alter the extent of the Mg^{2+} block of these channels¹²⁴. Another possibility is an upregulation of endogenous promoters of NMDA receptor function, such as arachidonic acid¹²⁵ and $InsP_3$ (ref. 126).

The nature of the retrograde messenger. The probable trigger for the induction of LTP is the entry of Ca^{2+} through NMDA channels located on the postsynaptic cell. But as we have seen, it is very likely that the potentiated response is maintained in part by presynaptic mechanisms. To reconcile these two observations, it was proposed that an intercellular signal is released from the postsynaptic site of induction to initiate increased transmitter release from the presynaptic terminal^{52,127}.

The first candidates to be considered were proteins. In addition to a tetanus-induced efflux of newly synthesised proteins from hippocampal slices¹²⁸ LTP is associated with an NMDA receptor-dependent increase in the protein content of hippocampal perfusates^{129,130}. The increases were slow to develop, apparently ruling out proteins as immediate retrograde messengers.

The next candidate to be examined was arachidonic acid. This unsaturated fatty acid satisfies several of the requirements for a retrograde messenger: (1) it is released from cultured neurons into the extracellular medium by the activation of NMDA receptors¹³¹, (2) there is an increase in its efflux¹³² and postsynaptic availability¹³³ following the induction of LTP, (3) inhibitors of phospholipase A_2 , an enzyme that liberates arachidonic acid from phospholipids, block the induction of LTP^{134–135}, and (4) the transient application of arachidonic acid to hippocampal synapses causes a slow-onset potentiation^{66,136}. Potential targets for arachidonic acid include not only the presynaptic terminal, where it may act to increase L-glutamate release⁶⁶, but also glial cells where it depresses L-glutamate uptake¹³⁷ and the postsynaptic cell, where, for example, it can potentiate NMDA receptor-mediated currents¹²⁵. Another phospholipase A_2 -derived lipid, platelet-activating factor, also has some of the properties expected of a retrograde messenger^{137,138}.

The possibility that nitric oxide (NO) may be a retrograde messenger in LTP has excited considerable interest. Like arachidonic acid, NO is released from cultured neurons exposed to NMDA¹³⁹. NO is derived from arginine in a reaction catalysed by NO synthase, and inhibitors of the enzyme have been reported to block the induction of LTP^{140–143}. Haemoglobin, a scavenger of NO which is presumably confined to the extracellular space, also blocks the induction of LTP^{141–143}, implying that NO (or another haem-binding molecule, such as CO) is released into the extracellular compartment. In addition, NO increases the frequency of miniature excitatory postsynaptic potentials (e.p.s.ps) in hippocampal cultures¹⁴². But although there is immunocytochemical evidence for NO synthase in hippocampal interneurons, there has been difficulty in obtaining evidence for its expression in pyramidal or granule cells¹⁴⁴. Furthermore, other laboratories¹⁴⁵, including our own, have not found a consistent block of LTP with NO synthase inhibitors. The story has been further complicated by the observation that under conditions where previous activation of the NMDA receptor system has disabled the induction mechanism^{28,29}, NO synthase inhibitors may promote the induction of LTP³⁰. Thus, the effect of NO synthase inhibitors may depend on the recent history of activity in the hippocampus. In summary, although NO remains an intriguing candidate, the evidence that it is a retrograde messenger is far from conclusive.

A general problem with the candidates discussed above is the time course of their action. The evidence for increased transmitter release is strongest for STP, that is from a few seconds to an hour or so after the inductive event. But inhibitors of arachidonic acid and NO synthesis both spare STP. Moreover, the potentiation produced by arachidonic acid is comparably slow to develop. Thus none of the proposed candidates has the properties expected of a *rapid* retrograde messenger. An alternative means of relaying postsynaptic activity is through alterations in activities of extracellular ions. One possibility is K^+ , which will be released from the postsynaptic cell during a tetanus to a degree that will, in part, reflect the level of activation of NMDA receptors. As discussed elsewhere¹⁴⁶, this could provide a signal to the presynaptic terminal through an interaction with presynaptic mGluRs¹⁴⁷, because the coupling of these receptors to PLC is strongly potentiated by extracellular K^+ (ref. 148).

Presynaptic modifications. Regulation of transmitter release could occur at any of the sequence of events leading from Ca^{2+} entry to exocytosis, through the mobilization, docking and fusion of vesicles at release sites in the presynaptic terminal.

BOX 4 LTP: Some unresolved issues

- (1) WHAT is the physiological significance of LTP? Specifically, is it a central component in the synaptic machinery of memory?
- (2) What percentage of excitatory synapses can be potentiated? Is LTP at an individual synapse a graded or an all-or-none event?
- (3) What are (1) the presynaptic, and (2) the postsynaptic mechanisms underlying expression of LTP? What is the relative contribution of these two components and how does this change with time?
- (4) How do changes in the number or structure of synapses contribute to LTP?
- (5) Do retrograde messengers exist? If so, what are they and how do they regulate neurotransmitter release?
- (6) How prevalent is NMDA receptor-independent LTP, and to what extent

- do the two forms of LTP share common mechanisms?
- (7) How do other neurotransmitter and neuromodulators, such as acetylcholine, monoamines and peptides, regulate the induction and expression of LTP?
- (8) Does LTP always decay or is there a non-decremental form in the brain? Can LTP be reversed (depotentiated)?
- (9) What is the extent and significance of long-term depression (LTD) in the hippocampus?
- (10) Can knowledge about the mechanisms of LTP be exploited to devise rational therapies for neurological disorders such as Alzheimer's disease?

LTP-related changes in Ca^{2+} homeostasis could in principle account for persistent changes in transmitter release. Ca^{2+} levels were found to be elevated in synaptosomes prepared from potentiated dentate gyrus 45 min after the induction of LTP¹⁴⁹, and this may explain the enhanced ability of potentiated synaptosomes to release preloaded transmitter. Another possibility is an increase in the size of the Ca^{2+} transient associated with each action potential, following the induction of LTP. The measurement of Ca^{2+} transients associated with single action potentials in hippocampal afferent terminals has not yet been reported; however, the Ca^{2+} signal produced in mossy fibre terminals by trains of stimuli is not changed following the induction of LTP in this NMDA receptor-independent pathway¹⁵⁰. Alternatively, LTP may be associated with an increase in the sensitivity to Ca^{2+} to one or more components of the release mechanism¹⁵¹. Because LTP is expressed as an enhanced response to single stimuli, it is processes controlling the rapid fusion of synaptic vesicles with release sites, and/or the formation of fusion pores, which are the most likely targets for regulation. Processes which govern the ability of the terminal to respond during sustained activity, such as the synthesis of transmitter, the transport and filling of vesicles and their release from the cytoskeletal cage, will contribute to LTP only to the extent that they influence either the probability of fusion, or the amount of transmitter packed into vesicles.

The nature of the retrograde messenger may give clues to the processes responsible for the sustained increase in transmitter release. Arachidonic acid stimulates basal phosphoinositide turnover in synaptosomes prepared from the dentate gyrus¹⁵², and, consistent with this finding, there is an increase in presynaptic phosphoinositide turnover in LTP¹⁴⁹. Arachidonic acid therefore could lead to an activation of presynaptic PKC both directly and as a consequence of the increased production of diacylglycerol. Among presynaptic substrates for PKC is the calmodulin-binding protein gap43, phosphorylation of which is increased in LTP^{153,154}. Because phosphorylated gap43 cannot bind calmodulin, it is possible that through the resulting increased availability of calmodulin, the phosphorylation of synaptic vesicle proteins such as CaMKII substrates synaptophysin and synapsin could be affected, leading to modulation of vesicle fusion and hence of transmitter release. The identity of presynaptic targets for NO, which could include guanylate cyclase and ADP ribosyltransferase¹⁴², have not been determined. Finally, a presynaptic mGluR could be coupled to transmitter release in a number of ways, as suggested by the coupling of mGluR1 to PI hydrolysis, arachidonic acid production, and cAMP levels¹⁵⁵.

Conclusion

The associative characteristics that define the induction criteria for NMDA receptor-dependent LTP have found an elegant and satisfying explanation in the voltage-dependent properties of the NMDA receptor/channel complex. In contrast, little is known about the biochemical cascades that are triggered by the permeation of Ca^{2+} through open NMDA channels and which

lead to the persistent enhancement of synaptic efficiency. The evidence considered here suggests that tetanus-induced potentiation proceeds in stages, beginning with a protein kinase-independent phase (STP), lasting less than ~1 hour, followed by three stages of LTP (LTP1–3), requiring protein phosphorylation, protein synthesis from existing mRNAs, and gene transcription, respectively. The expression of synaptic potentiation probably involves both pre- and postsynaptic mechanisms, not necessarily in the same proportion at each stage, the one leading to an increase in transmitter release and the other to an increase in the number or change in the properties of the ion channels which mediate synaptic transmission. Activity-induced changes in the morphology or number of spines may also contribute to changes in synaptic efficiency, as suggested by a number of electron-microscopic studies^{156,157}. Advances in microscopy may soon allow the real-time visualization of any such changes¹⁵⁸.

In this review we have charted the substantial progress which has been made in understanding the cellular and molecular basis of NMDA receptor-dependent LTP in the hippocampus. It is part of the fascination of LTP that it can be studied experimentally at many levels, from the molecular to the behavioural; at the same time, knowledge about properties of LTP feeds directly into theoretical investigations of information storage in distributed neural networks. This catholicity of interest is reflected in the scope of the many questions that remain (Box 4). In the end, the overriding motivation for studying synaptic plasticity in the brain is the hope of gaining an understanding of the physical basis of memory in health and disease, and it is the nature of the link between LTP and memory that is likely to provide a major focus for research in the future. □

T. V. P. Bliss is at the Division of Neurophysiology and Neuropharmacology, National Institute of Medical Research, Mill Hill, London NW7 1AA, UK; G. L. Collingridge is at the Department of Pharmacology, The University of Birmingham, Birmingham B15 2TT, UK.

1. Hebb, D. O. *The Organization of Behaviour* (Wiley, New York, 1949).
2. Konorski, J. *Conditioned Reflexes and Neuron Organisation* (Cambridge Univ. Press, Cambridge, 1948).
3. Bliss, T. V. P. & Lømo, T. *J. Physiol., Lond.* **232**, 331–356 (1973).
4. Bliss, T. V. P. & Gardner-Medwin, A. R. *J. Physiol., Lond.* **232**, 357–374 (1973).
5. Morris, R. G. M., Davis, S. & Butcher, S. P. *Phil. Trans. R. Soc.* **329**, 187–204 (1990).
6. Doyere, V. & Laroche, S. *Hippocampus* **2**, 39–48 (1992).
7. Larson, J., Wong, D. & Lynch, G. *Brain Res.* **368**, 347–350 (1986).
8. Rose, G. M. & Dunwiddie, T. V. *Neurosci. Lett.* **69**, 244–248 (1986).
9. Otto, T., Eichenbaum, H., Wiener, S. I. & Wible, C. G. *Hippocampus* **1**, 181–192 (1991).
10. McNaughton, B. L., Douglas, R. M. & Goddard, G. V. *Brain Res.* **157**, 277–294 (1978).
11. Lovinger, D. M. & Routtenberg, A. *J. Physiol., Lond.* **400**, 321–334 (1988).
12. Malenka, R. C. *Neuron* **6**, 53–60 (1991).
13. Levy, W. B. & Steward, O. *Brain Res.* **175**, 233–245 (1979).
14. Andersen, P., Sundberg, S. H., Sveen, O. & Wigström, H. *Nature* **266**, 736–737 (1977).
15. Lynch, G., Dunwiddie, T. & Gribkoff, V. *Nature* **266**, 737–739 (1977).
16. Kelso, S. R., Ganong, A. H. & Brown, T. H. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5326–5330 (1986).
17. Wigström, H., Gustafsson, B., Huang, Y.-Y. & Abrahams, W. C. *Acta physiol. scand.* **126**, 317–319 (1986).
18. Sastry, B. R., Goh, J. W. & Auyeung, A. *Science* **232**, 988–990 (1986).
19. Malinow, R. *Science* **252**, 722–724 (1991).
20. Malinow, R. & Miller, J. P. *Nature* **320**, 529–530 (1986).
21. Ascher, P. & Nowak, L. *J. Physiol., Lond.* **399**, 247–266 (1988).
22. Collingridge, G. L., Herron, C. E. & Lester, R. A. J. *J. Physiol., Lond.* **399**, 283–300 (1988).
23. Davies, C. H., Starkey, S. J., Pozza, M. F. & Collingridge, G. L. *Nature* **349**, 609–611 (1991).
24. Collingridge, G. L., Kehil, S. J. & McLennan, H. *J. Physiol., Lond.* **334**, 33–46 (1983).

25. Coan, E. J., Saywood, W. & Collingridge, G. L. *Neurosci. Lett.* **80**, 111–114 (1987).
26. Bashir, Z. I., Tam, B. & Collingridge, G. L. *Neurosci. Lett.* **108**, 261–266 (1990).
27. Kauer, J. A., Malenka, R. C. & Nicoll, R. A. *Nature* **334**, 250–252 (1988).
28. Coan, E. J., Irving, A. J. & Collingridge, G. L. *Neurosci. Lett.* **105**, 205–210 (1989).
29. Huang, Y.-Y., Colino, A., Selig, D. K. & Malenka, R. C. *Science* **255**, 730–733 (1992).
30. Izumi, Y., Clifford, D. B. & Zorumski, C. F. *Science* **257**, 1273–1276 (1992).
31. Reymann, K. G. & Matthies, H. *Neurosci. Lett.* **98**, 166–171 (1989).
32. Izumi, Y., Clifford, D. B. & Zorumski, C. F. *Neurosci. Lett.* **122**, 187–190 (1991).
33. McGuinness, N., Anwyl, R. & Rowan, M. *Eur. J. Pharmacol.* **197**, 231–232 (1991).
34. Collingridge, G. L. *et al. Int. Acad. Biomed. Drug Res.* **2**, 41–49 (1991).
35. Radpour, S. & Thomson, A. M. *Neurosci. Lett.* **138**, 119–122 (1992).
36. Bortolotto, Z. A. & Collingridge, G. L. *Neuropharmacology* **32**, 1–9 (1993).
37. Lynch, G., Larson, J., Kelso, S., Barriounevo, G. & Schottler, F. *Nature* **305**, 719–721 (1983).
38. MacDermott, A. B., Mayer, M. L., Westbrook, G. L., Smith, S. J. & Barker, J. L. *Nature* **321**, 519–522 (1986).
39. Jahr, C. E. & Stevens, C. F. *Nature* **325**, 522–525 (1987).
40. Guthrie, P. B., Segal, M. & Kater, S. B. *Nature* **354**, 76–80 (1991).
41. Regehr, W. G. & Tank, D. W. *Nature* **345**, 807–810 (1990).
42. Müller, W. & Connor, J. A. *Nature* **354**, 73–76 (1991).
43. Malenka, R. C., Lancaster, B. & Zucker, R. S. *Neuron* **9**, 121–128 (1992).
44. Alford, S. & Collingridge, G. L. in *Excitatory Amino Acids and Second Messenger Systems* (eds Teichberg, Turski, V. I.) 43–53 (Springer, Berlin, 1992).
45. Obenaus, A., Mody, I. & Baimbridge, K. G. *Neurosci. Lett.* **98**, 172–178 (1989).
46. Harvey, J. & Collingridge, G. L. *Neurosci. Lett.* **139**, 197–200 (1992).
47. Gasic, G. P. & Hollmann, M. A. *Rev. Physiol.* **54**, 507–536 (1992).
48. Malenka, R. C., Kauer, J. A., Zucker, R. S. & Nicoll, R. A. *Science* **242**, 81–84 (1988).
49. Malenka, R. C., Kauer, J. A., Perkel, D. J. & Nicoll, R. A. *Trends Neurosci.* **12**, 444–450 (1989).
50. Skrede, K. & Mølle-Sørensen, D. *Brain Res.* **208**, 436–441 (1981).
51. Dolphin, A. C., Errington, M. L. & Bliss, T. V. P. *Nature* **297**, 496–498 (1982).
52. Bliss, T. V. P., Douglas, R. M., Errington, M. L. & Lynch, M. A. *J. Physiol., Lond.* **377**, 391–408 (1986).
53. Aniksztejn, L., Roisin, M. P., Amselem, R. & Ben-Ari, Y. *Neuroscience* **28**, 387–392 (1989).
54. Bliss, T. V. P., Errington, M. L. & Lynch, M. A. *Adv. exp. med. Biol.* **268**, 269–278 (1990).
55. Feasey, K. J., Lynch, M. A. & Bliss, T. V. P. *Brain Res.* **364**, 39–44 (1986).
56. McNaughton, B. L. *J. Physiol., Lond.* **324**, 249–262 (1982).
57. Kauer, J. A., Malenka, R. C. & Nicoll, R. A. *Neuron* **1**, 911–917 (1988).
58. Müller, D. & Lynch, G. *Proc. natn. Acad. Sci. U.S.A.* **85**, 9346–9350 (1988).
59. Bashir, Z. I., Alford, S., Davies, S. N., Randall, A. D. & Collingridge, G. L. *Nature* **349**, 156–158 (1991).
60. Berretta, N. *et al. Eur. J. Neurosci.* **3**, 850–854 (1991).
61. Xie, X., Berger, T. W. & Barriounevo, G. *J. Neurophysiol.* **67**, 1009–1013 (1992).
62. Asztely, F., Wigström, H. & Gustafsson, B. *Eur. J. Neurosci.* **4**, 681–690 (1992).
63. Lynch, G., Gribkoff, V. & Deadwyler, S. A. *Nature* **263**, 151–153 (1976).
64. Taube, J. S. & Schwartzkroin, P. A. *J. Neurosci.* **8**, 1632–1644 (1988).
65. Davies, S. N., Lester, R. A. J., Reymann, K. G. & Collingridge, G. L. *Nature* **338**, 500–503 (1989).
66. Williams, J. H., Errington, M. L., Lynch, M. A. & Bliss, T. V. P. *Nature* **341**, 739–742 (1989).
67. Voronin, L. L. *Neuroscience* **10**, 1051–1069 (1983).
68. Bekkers, J. M. & Stevens, C. F. *Nature* **346**, 724–729 (1990).
69. Malinow, R. & Tsien, R. W. *Nature* **346**, 177–180 (1990).
70. Voronin, L. L., Kuhnt, U. & Gusev, A. G. *Exp. Brain Res.* **89**, 288–299 (1992).
71. Foster, T. C. & McNaughton, B. L. *Hippocampus* **1**, 79–91 (1991).
72. Kullmann, D. M. & Nicoll, R. A. *Nature* **357**, 240–244 (1992).
73. Larkman, A., Hannay, T., Stratford, K. & Jack, J. *Nature* **360**, 70–73 (1992).
74. Manabe, T., Renner, P. & Nicoll, R. A. *Nature* **355**, 50–55 (1992).
75. Malgaroli, A. & Tsien, R. W. *Nature* **357**, 134–139 (1992).
76. Oliver, M. W., Baudry, M. & Lynch, G. *Brain Res.* **505**, 233–238 (1989).
77. Halpain, S. & Greengard, P. *Neuron* **5**, 237–246 (1990).
78. Bär, P. R., Wiegant, F., Lopes da Silva, F. H. & Gispén, W. H. *Brain Res.* **321**, 381–385 (1984).
79. Akers, R., Lovinger, D., Colley, P., Linden, D. & Routtenberg, A. *Science* **231**, 587–589 (1986).
80. Klann, E., Chen, S.-J. & Sweatt, J. D. *J. Biol. Chem.* **266**, 24253–24256 (1991).
81. Lovinger, D. M., Wong, K. L., Murakami, K. & Routtenberg, A. *Brain Res.* **436**, 177–183 (1987).
82. Reymann, K. G., Frey, U., Jork, R. & Matthies, H. *Brain Res.* **440**, 305–314 (1988).
83. Malinow, R., Madison, D. V. & Tsien, R. W. *Nature* **335**, 821–824 (1988).
84. Malenka, R. C. *et al. Nature* **340**, 554–557 (1989).
85. Reymann, K. G., Davies, S. N., Matthies, H., Kase, H. & Collingridge, G. L. *Eur. J. Neurosci.* **2**, 481–486 (1990).
86. Wang, J. & Feng, D.-P. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2576–2580 (1992).
87. Müller, D., Buchs, P.-A., Dunant, Y. & Lynch, G. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4073–4077 (1990).
88. Malinow, R., Schulman, H. & Tsien, R. W. *Science* **245**, 862–866 (1989).
89. Huang, Y.-Y., Colley, P. A. & Routtenberg, A. *Neuroscience* **49**, 819–827 (1992).
90. Hu, G.-Y. *et al. Nature* **328**, 426–429 (1987).
91. Malenka, R. C., Madison, D. V. & Nicoll, R. A. *Nature* **321**, 175–177 (1986).
92. Müller, D., Turnbull, J., Baudry, M. & Lynch, G. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6997–7000 (1988).
93. Gustafsson, B., Huang, Y.-Y. & Wigström, H. *Neurosci. Lett.* **85**, 77–81 (1988).
94. Mody, I., Baimbridge, K. G. & Miller, J. J. *Neuropharmacology* **23**, 625–631 (1984).
95. Reymann, K. G., Brodemann, R., Kase, H. & Matthies, H. *Brain Res.* **461**, 388–392 (1988).
96. Ito, I., Hidaka, H. & Sugiyama, H. *Neurosci. Lett.* **121**, 119–121 (1991).
97. Silva, A. J., Stevens, C. F., Tonegawa, S. & Wang, Y. *Science* **257**, 201–206 (1992).
98. Lisman, J. E. & Goldring, M. A. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5320–5324 (1988).
99. Molloy, S. S. & Kennedy, M. B. *Proc. natn. Acad. Sci. U.S.A.* **88**, 4756–4760 (1991).
100. Chetkovich, D. M., Gray, R., Johnston, D. & Sweatt, J. D. *Proc. natn. Acad. Sci. U.S.A.* **88**, 6467–6471 (1991).
101. O'Dell, T. J., Kandel, E. R. & Grant, S. G. N. *Nature* **353**, 558–560 (1991).
102. Bading, H. & Greengard, M. E. *Science* **253**, 912–914 (1991).
103. Krug, M., Lossner, B. & Ott, T. *Brain Res. Bull.* **13**, 39–42 (1984).
104. Otani, S., Marshall, C. J., Tate, W. P., Goddard, G. V. & Abraham, W. C. *Neuroscience* **28**, 519–526 (1989).
105. Frey, U., Krug, M., Brödemann, R., Reymann, K. & Matthies, H. *Neurosci. Lett.* **97**, 135–139 (1989).
106. Fazeli, M. S., Corbett, J., Dunn, M. J., Dolphin, A. C. & Bliss, T. V. P. *J. Neurosci.* **13**(4) (in the press).
107. Fazeli, M. S., Errington, M. L., Dolphin, A. C. & Bliss, T. V. P. *Brain Res.* **521**, 247–253 (1990).
108. Abraham, W. C., Dragunow, M. & Tate, W. P. *Molec. Neurobiol.* **5**, 297–314 (1991).
109. Cole, A. J., Saffen, D. W., Baraban, J. M. & Worley, P. F. *Nature* **340**, 474–476 (1989).
110. Wisden, W. *et al. Neuron* **4**, 603–614 (1990).
111. Dragunow, M. *et al. Neurosci. Lett.* **101**, 274–280 (1989).
112. Nikolaev, E., Tischmeyer, W., Krug, M., Matthies, H. & Kaczmarek, L. *Brain Res.* **560**, 346–349 (1991).
113. Mackler, S. A., Brooks, B. P. & Eberwine, J. H. *Neuron* **9**, 539–548 (1992).
114. Greengard, P., Jen, J., Nairn, A. C. & Stevens, C. F. *Science* **253**, 1135–1138 (1991).
115. Wang, L.-Y., Salter, M. W. & MacDonald, J. F. *Science* **253**, 1132–1135 (1991).
116. Sommer, B. *et al. Science* **249**, 1580–1585 (1990).
117. Sommer, B., Kohler, M., Sprengel, R. & Seeburg, P. H. *Cell* **67**, 11–19 (1992).
118. Ito, I., Tanabe, S., Khoda, A. & Sugiyama, H. *J. Physiol., Lond.* **424**, 533–543 (1990).
119. Isaacson, J. S. & Nicoll, R. A. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10936–10940 (1991).
120. Tang, C. M., Shi, Q. Y., Katchman, A. & Lynch, G. *Science* **254**, 288–290 (1991).
121. Staubli, U., Kessler, M. & Lynch, G. *Psychobiology* **18**, 377–381 (1990).
122. Asztely, F., Hanse, E., Wigström, H. & Gustafsson, B. *Synapse* **11**, 342–345 (1992).
123. Kelso, S. R., Nelson, T. E. & Leonard, J. P. *J. Physiol., Lond.* **449**, 705–718 (1992).
124. Chen, L. & Huang, L.-Y. *Nature* **356**, 521–523 (1992).
125. Miller, B., Sarantis, M., Traynelis, S. F. & Attwell, D. *Nature* **355**, 722–725 (1992).
126. Markram, H. & Segal, M. *J. Physiol., Lond.* **447**, 513–533 (1992).
127. Bliss, T. V. P. & Dolphin, A. C. in *The Neurobiology of Learning and Memory* (eds McGaugh, J. L. & Lynch, G.) (Guilford, New York, 1984).
128. Duffy, C., Teyler, T. J. & Shashoua, V. E. *Science* **212**, 1148–1151 (1981).
129. Fazeli, M. S., Errington, M. L., Dolphin, A. C. & Bliss, T. V. P. *Brain Res.* **473**, 51–59 (1988).
130. Otani, S., Roisin-Lallemant, M.-P. & Ben-Ari, Y. *Neuroscience* **47**, 265–272 (1992).
131. Dumuis, A., Sebben, M., Haynes, L., Pin, J.-P. & Bockaert, J. *Nature* **336**, 68–70 (1988).
132. Bliss, T. V. P., Errington, M. L., Lynch, M. A. & Williams, J. H. *Cold Spring Harb. Symp. quant. Biol.* **55**, 119–129 (1990).
133. Clements, M. P., Bliss, T. V. P. & Lynch, M. A. *Neuroscience* **45**, 379–389 (1991).
134. Okada, D., Yamagishi, S. & Sugiyama, H. *Neurosci. Lett.* **100**, 141–146 (1989).
135. Williams, J. H. & Bliss, T. V. P. *Neurosci. Lett.* **107**, 301–306 (1989).
136. Barbour, B., Szatkowski, M., Ingledew, N. & Attwell, D. *Nature* **342**, 918–920 (1989).
137. Arai, A. & Lynch, G. *Eur. J. Neurosci.* **4**, 411–419 (1992).
138. Clark, G. D., Happel, L. T., Zorumski, C. F. & Bazan, N. G. *Neuron* **9**, 1211–1216 (1992).
139. Garthwaite, J., Charles, S. L. & Chess-Williams, R. *Nature* **336**, 385–388 (1988).
140. Böhm, G. A., Bon, C., Stutzmann, J.-M., Doole, A. & Blanchard, J.-C. *Fur. J. Pharmacol.* **199**, 379–381 (1991).
141. Schuman, E. M. & Madison, D. V. *Science* **254**, 1503–1506 (1991).
142. O'Dell, T. J., Hawkins, R. D., Kandel, E. R. & Arancio, O. *Proc. natn. Acad. Sci. U.S.A.* **88**, 11285–11289 (1991).
143. Haley, J. E., Wilcox, G. L. & Chapman, P. F. *Neuron* **8**, 211–216 (1992).
144. Brett, S. B. & Snyder, S. H. *Neuron* **8**, 3–11 (1992).
145. Gribkoff, V. K. & Lum-Ragan, J. T. *J. Neurophysiol.* **68**, 639–642 (1992).
146. Collingridge, G. L. *Exp. Physiol.* **77**, 771–797 (1992).
147. Baskys, A. & Malenka, R. C. *J. Physiol., Lond.* **444**, 687–701 (1991).
148. Irving, A. J., Collingridge, G. L. & Schofield, J. G. *Cell Calcium* **13**, 293–301 (1992).
149. Lynch, M. A. & Voss, K. L. *J. Neurochem.* **56**, 113–118 (1991).
150. Regehr, W. G. & Tank, D. W. *Neuron* **7**, 451–459 (1991).
151. Lynch, M. A. & Bliss, T. V. P. *Brain Res.* **369**, 405–408 (1986).
152. Lynch, M. A. & Voss, K. L. *J. Neurochem.* **55**, 215–221 (1990).
153. Linden, D. J., Wong, K. L., Sheu, F.-S. & Routtenberg, A. *Brain Res.* **458**, 142–146 (1988).
154. Giannotti, C., Nunzi, M. G., Gispén, W. H. & Corradetti, R. *Neuron* **8**, 843–848 (1992).
155. Aramori, I. & Nakanishi, S. *Neuron* **8**, 757–765 (1992).
156. Desmond, N. L. & Levy, W. B. *Synapse* **5**, 139–143 (1990).
157. Geinisman, Y., de Toledo-Morrell, L. & Morrell, F. *Brain Res.* **566**, 77–88 (1991).
158. Hosokawa, T., Bliss, T. V. P. & Fine, A. *Neuroreport* **3**, 477–480 (1992).
159. Jeffery, K. J., Abraham, W. C., Dragunow, M. & Mason, S. E. *Molec. Brain Res.* **8**, 267–274 (1990).
160. Andersen, P., Sundberg, S. H., Sveen, O., Swann, J. W. & Wigström, H. *J. Physiol., Lond.* **302**, 463–482 (1980).
161. Abraham, W. C., Bliss, T. V. P. & Goddard, G. V. *J. Physiol., Lond.* **363**, 335–349 (1985).
162. Bonhoeffer, T., Staiger, V. & Aertsen, A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8113–8117 (1989).
163. Harris, E. W. & Cotman, C. W. *Neurosci. Lett.* **70**, 132–137 (1986).
164. Johnston, D., Williams, S., Jaffe, D. & Gray, R. A. *Rev. Physiol.* **54**, 489–505 (1992).
165. Grover, L. M. & Teyler, T. J. *Nature* **347**, 477–479 (1990).
166. Turner, R. W., Baimbridge, K. G. & Miller, J. J. *Neuroscience* **7**, 1411–1416 (1982).
167. Aniksztejn, L. & Ben-Ari, Y. *Nature* **349**, 67–69 (1991).
168. Publicover, S. J. *Exp. Brain Res.* **84**, 680–684 (1991).
169. Davies, S. N. & Collingridge, G. L. *Proc. R. Soc. B* **236**, 373–384 (1989).
170. Andreasen, M., Lambert, J. D. C. & Jensen, M. S. *J. Physiol., Lond.* **414**, 317–336 (1989).
171. Dale, N. & Roberts, A. *J. Physiol., Lond.* **363**, 35–59 (1985).
172. Lester, R. A. J., Clements, J. D., Westbrook, G. L. & Jahr, C. E. *Nature* **346**, 565–567 (1990).
173. Collingridge, G. L., Herron, C. E. & Lester, R. A. J. *J. Physiol., Lond.* **399**, 301–312 (1988).
174. Herrero, I., Miras-Portugal, M. T. & Sánchez-Prieto, J. *Nature* **360**, 163–166 (1992).

ACKNOWLEDGEMENTS. We thank J. Williams and M. Errington, and S. Alford and B. Frenguelli for providing the data illustrated in Figs 1 and 2, respectively. J. Satchell for help with the diagrams in Boxes 2 and 3, and we thank other colleagues in our two laboratories and fellow members of the Human Frontiers Science Program project grant on 'Properties and mechanisms of LTP and LTD in the hippocampus' for valuable discussions.

The 1908 Tunguska explosion: atmospheric disruption of a stony asteroid

Christopher F. Chyba^{*†}, Paul J. Thomas[‡] & Kevin J. Zahnle^{*}

^{*} Space Science Division, MS 245-3, NASA Ames Research Center, Moffett Field, California 94035, USA

[‡] Department of Physics and Astronomy, University of Wisconsin, Eau Claire, Wisconsin 54702, USA

The explosion over Tunguska, Central Siberia, in 1908 released 10 to 20 megatons (high explosive equivalent) of energy at an altitude of about 10 km. This event represents a typical fate for stony asteroids tens of metres in radius entering the Earth's atmosphere at common hypersonic velocities. Comets and carbonaceous asteroids of the appropriate energy disrupt too high, whereas typical iron objects reach and crater the terrestrial surface.

THE explosion¹⁻⁵ on 30 June 1908 over Tunguska has inspired many exotic explanations. Antimatter⁶, a small black hole⁷ and, inevitably, an exploding flying saucer², have all been proposed as means of liberating tens of megatons of energy in the atmosphere without cratering the Earth's surface. Quantitative explanations in terms of a less exotic object have often suggested that it must have been extremely underdense (with an effective density $\rho_m \approx 10^{-3}$ – 10^{-2} g cm⁻³) to have exploded before reaching the ground⁸⁻¹². Contrary to that claim, we show that the Tunguska explosion is a typical fate for stony asteroids with radii ~ 30 m entering the Earth's atmosphere at common hypersonic velocities. Short-period and long-period comets with appropriate energies explode far too high in the atmosphere to account for the observations, whereas iron objects (with rare exceptions) explode too low, or not at all. Our model is also consistent with the fate of the Revelstoke object, observed¹³ to have catastrophically exploded in the atmosphere with an energy of tens of kilotons⁴. These results affect assessments of the hazard posed by impacts of small comets and asteroids.

Characteristics of the Tunguska event

The Tunguska event has been variously estimated as liberating between 4×10^{23} (ref. 14) and 4×10^{25} erg, (ref. 12) or 10–10³ megatons (Mton). The best energy estimates are based on air and seismic wave records, compared with nuclear airbursts of comparable yield. Hunt *et al.*¹⁴ thereby find an explosive energy of 10 ± 5 Mton; a more recent analysis by Ben-Menahem¹⁵ finds 12.5 ± 2.5 Mton. Similar estimates derive from observed forest destruction, scaled from the uprooting of trees in nuclear weapons tests^{14,16}.

Seismic waves were excited by a vertical point impulse of 7×10^{18} dyn s (ref. 15). The impulse has been interpreted¹² as the initial vertical momentum of the impactor, which implies a bolide kinetic energy of $\sim 4 \times 10^{25}$ erg. This interpretation is problematic, as the impulse striking the ground is largely from the blast wave generated by the explosion, and not due to the detached ballistic shock from the decelerated impactor¹⁵. Turco *et al.*¹² distinguish between the Tunguska terminal explosion and energy released by ablation and deceleration during atmospheric entry, with the explosion representing just the last 1% of the bolide's energy. This requires the bolide to be extremely underdense ($\rho_m < 0.01$ g cm⁻³) by the time it enters the upper atmosphere.

A reported¹⁷ iridium spike coincident with Tunguska in Antarctic ice corresponds to an impactor with energy $\geq 10^{25}$ erg. New analyses on Antarctic samples, however, have found no detectable iridium imprint above background due to cosmic dust, contradicting the earlier result¹⁸. No evidence for pre-

dicted¹² elevated nitrate production has been found in a Greenland ice core¹⁹.

Contrary to interpretations of the Tunguska object as a friable, underdense comet, Sekanina²⁰ favours a bolide with a very high strength, arguing that the object would explode instantly on disruption. Levin and Bronshten²¹ reach similar conclusions, by analogy with typical terminal-flare meteors. Shoemaker⁴ asserts that Tunguska-like explosions should be a common fate of small bolides entering the terrestrial atmosphere, provided only that they are less strong than iron objects.

The terminal explosion over Tunguska is typically estimated to have occurred at roughly an atmospheric scale height $H \approx 8.4$ km. Ben-Menahem¹⁵ has compared differing arrival times for Rayleigh and SH body waves recorded at Irkutsk; he finds an explosion height of 8.5 km. This agrees with simulations^{16,22} of the treefall pattern, which suggests explosion altitudes of 5–10 km.

The inclination angle θ (measured from the horizontal) of the Tunguska object's trajectory has been controversial. Attempts^{3,20,21} to determine θ from eyewitness reports typically give shallow trajectories, in the range 5–17°. Such low inclinations are inconsistent with attempts to simulate the treefall pattern at the Tunguska site. Zotkin and Tsikulin²³ measured the azimuths of 40,000 felled tree trunks over 2,200 km² at the Tunguska site, and reproduced the observed 'butterfly' pattern in the laboratory by superimposing a terminal point charge and an inclined line charge. Their investigations require $10^\circ < \theta < 60^\circ$, with a preferred value of 30°. Korobeinikov *et al.*²² simulated the treefall pattern using a numerical shock model, finding that observations can be matched only for $30^\circ < \theta < 45^\circ$, with a preferred value of 40°. They further assert that this value is consistent with eyewitness accounts.

Evidently it is difficult to determine the true inclination of the Tunguska bolide, and eyewitness reports gathered 20 years after the event are particularly suspect. In this investigation, we usually take θ to be the most probable entry angle for an incident object, 45°, but consider a range of other possible values.

Atmospheric entry of the bolide

A cosmic object entering the atmosphere loses its kinetic energy through deceleration and ablation²⁴⁻²⁶. Deceleration can be described by the equation

$$m \frac{dv}{dt} = -\frac{1}{2} C_D \rho_a A v^2 + \frac{g}{m} \sin \theta \quad (1)$$

where r is the bolide's radius, A its cross-sectional area, ρ_a atmospheric density, g gravitational acceleration (a function of height), t time, and C_D drag coefficient. The object's surface is heated by radiation from the atmospheric shock front. This heat is shed efficiently by ablation. The resulting change in mass is

[†] Present address: Laboratory for Extraterrestrial Physics, Code 693, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA.

given by

$$Q \frac{dm}{dt} = -\frac{1}{2} C_H \rho_a A v^3 \quad (2)$$

where Q is the heat of ablation and C_H is the heat transfer coefficient. Q is a function of material type and the specific process of ablation. To derive the values of Q used here, we begin with the heat of vaporization appropriate to iron and stony meteorites²⁴, 8×10^{10} erg g⁻¹. We then use observed ablation coefficients for cometary, carbonaceous and stony meteors to calculate Q for comets and carbonaceous asteroids, following Chyba *et al.*²⁷ (see Table 1).

Observation indicates²⁵ that C_H is effectively constant at $C_H \approx 0.1$ above ~ 30 km. This altitude range includes most visible meteors. But as a bolide descends to lower altitudes, C_H declines inversely with atmospheric density, so that $\dot{m}$, which had been increasing, becomes effectively constant. This upper limit on $\dot{m}$ occurs because large objects ablate mainly by absorbing thermal radiation emitted by the hot, shocked gases concentrated in front of the impactor. The temperature attained by the shocked gas is strongly regulated by thermal ionization^{28,29} to $\sim 25,000$ – $30,000$ K, with weak dependence on the velocity, size and composition of the impactor. The maximum ablation rate is therefore $Q\dot{m} \approx A\sigma T^4$. These considerations suggest rewriting equation (2) as²⁹

$$Q \frac{dm}{dt} = -A \min \left(\frac{1}{2} C_H^0 \rho_a v^3, \sigma T^4 \right) \quad (3)$$

where $C_H^0 = 0.1$ and $T = 25,000$ K. This parametrization roughly reproduces the results of ref. 30.

The bolide's trajectory angle θ varies as

$$\frac{d\theta}{dt} = \frac{g \cos \theta}{v} - \frac{C_L \rho_a A v}{2m} - \frac{v \cos \theta}{R_\oplus + h} \quad (4)$$

Here $R_\oplus$ is the Earth's radius, h the bolide's height above the

terrestrial surface, and C_L the lift coefficient. Passey and Melosh³¹ suggest taking $C_L \approx 10^{-3}$, on the basis of their investigations of crater fields. We set $C_L = 10^{-3}$. Even for trajectories as low as $\theta = 15^\circ$, varying C_L over the range 10^{-3} – 1 changes the altitude at which a Tunguska-sized stony asteroid airbursts by only $\sim 1\%$.

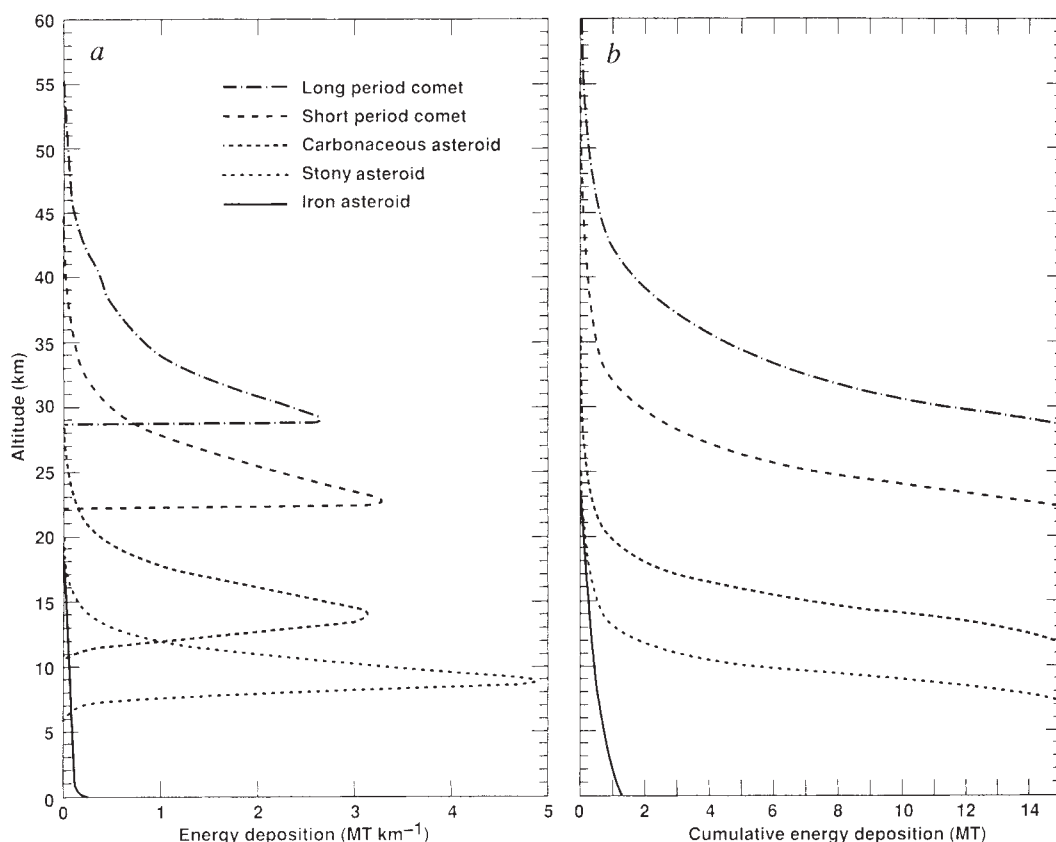
Equations (1) and (2) can be solved analytically for spherical impactors in an isothermal atmosphere if θ and C_H are held constant²⁵. The analytical solution does not, however, allow for objects breaking up in response to aerodynamic forces. We therefore apply a finite-difference approach to solving the required equations.

Catastrophic fragmentation

Catastrophic fragmentation is a likely means of producing an atmospheric explosion of a bolide. By spreading the impactor's mass over a wide area, fragmentation increases the amount of atmosphere intercepted and so enhances ablation and aerobraking; hence a fragmenting object stops more abruptly, surrendering its kinetic energy more explosively, than does a non-fragmenting object^{21,31,32}.

Objects much greater than ~ 1 km in diameter do not fragment while traversing the Earth's atmosphere, as an atmospherically induced pressure wave has insufficient time to cross the object before impact. In effect, large objects moving at hypersonic velocities do not have time to 'see' the atmosphere before cratering the surface²⁷. (This is consistent with the existence of the 14-km-wide Lappajärvi crater in Finland, apparently the result of an impactor of carbonaceous chondritic composition³³; this crater would have been excavated by a carbonaceous asteroid ~ 1 km in diameter.) Sufficiently small objects will either be entirely ablated, or be aerobraked to free-fall speeds. Atmospheric entry of objects in the size range ~ 10 – 100 m, however, is dominated by fragmentation, although the precise size range depends strongly on object type and velocity.

FIG. 1 Airburst altitudes for five 15-Mton candidate Tunguska bolides incident at 45° . Comets and carbonaceous asteroids deposit their energy too high in the atmosphere to account for the Tunguska explosion, whereas iron objects reach and crater the surface. Typical stony asteroids, however, deliver the bulk of their energy near an altitude of 9 km, and this is consistent with observations.



Deformation and fragmentation occur because of differential atmospheric pressure across the object. The leading face of the impactor is subjected to an average pressure $p_s \approx C_D \rho_a v^2/2$, whereas the pressure on the trailing face is much smaller. Integrated over the surface of the impactor, this difference produces the drag force of equation (1). Pressures on the side of the objects are also much smaller³⁴ than p_s , so that the object is essentially not laterally confined. Impactors fragment as the result of this aerodynamic stress^{24,25,31,35,36}. Fragmentation occurs when p_s exceeds a characteristic strength of the material. Because the details of bolide failure are poorly understood²⁴, we select characteristic strengths for various bolides using the following considerations.

Fragmentation of 'Sun-grazing' and 'Jupiter-grazing' comets, presumably due to tidal stresses, suggests³⁷ that at least some comets may have extremely low tensile strengths of 10^3 – 10^5 dyn cm⁻². Objects have fragmented in the terrestrial atmosphere with strengths²⁵ as low as 6×10^5 dyn cm⁻². A typical strength of a chondritic impactor is 1×10^7 to 5×10^7 dyn cm⁻², although stronger stony objects can have strengths²⁵ as high as 2×10^7 – 5×10^8 dyn cm⁻². By contrast iron bodies are strong and usually do not fragment until they penetrate to ≤ 10 km of the ground. The effective strength for an iron impactor²⁵ is 4×10^8 – 2×10^9 dyn cm⁻². For an object entering the atmosphere at 20 km s^{-1} , aerodynamic stresses reach $\sim 6 \times 10^8$ dyn cm⁻² two scale heights up, exceeding most of the strengths cited above.

A model for bolide deformation

A right circular cylinder moving along its axis of symmetry makes a highly idealized but straightforward model of a deforming impactor. To represent a roughly equidimensional bolide, we choose a 'cubical' cylinder, taking its height h and diameter $2r$ to be initially equal. We now consider the forces on the cylinder as it passes through the atmosphere. The front face sees a pressure $p_s = C_D \rho_a v^2/2$. For a cylinder³⁸, $C_D \approx 1.7$. Provided that the sound travel time h/c (c is the sound speed in the object) is short compared with the time $H \csc \theta/v$ for the impactor to fall through a scale height, stresses parallel to the axis should roughly be in hydrostatic equilibrium, with the axial stress at any point within the object being that required to decelerate the trailing mass. Thus the axial stress decreases linearly from its peak value p_s at the front face to some low value $\ll p_s$ at the rear face.

The air pressure against the side walls is generally much less³⁴ than p_s and except near the back of the cylinder is small compared with the axial stress. As the cylinder descends into deeper air, axial stress increases until elastic failure occurs. The cylinder flows outward, transversely to the direction of motion. Because the driving pressure p_s rises exponentially as the bolide descends through the atmosphere, its effective cross-section increases exponentially with time. A real disrupting cylinder would fail first at its leading edge, where pressures are highest. The locus of failure thereafter moves backwards into the cylinder. Because the pressures are always highest at the front the expansion is

always fastest there, although the most rapidly expanding elements may be swept away by the flow.

As a first approximation, we assume that the cylinder deforms globally to become a squatter version of itself. The average interior pressure is $p_s/2$. Neglecting the confining air pressure against the side walls, and assuming that the material strength of the cylinder has been exceeded, a global approximation to the force balance on the side walls is

$$(2\pi rh)(\frac{1}{4}C_D\rho_a v^2) \approx m \frac{d^2 r}{dt^2} \quad (5)$$

where the inertial mass is identified with the mass m of the cylinder. Assuming that the density ρ_m of the disrupted cylinder remains constant, h and m can be eliminated from equation (5), to give

$$r \frac{d^2 r}{dt^2} = \frac{C_D \rho_a v^2}{2\rho_m} \quad (6)$$

This equation is functionally identical (differing by only a factor of order unity) to the analogous equation derived by Zahnle²⁹ using a wholly different argument. Our finite-difference scheme solves equations (1)–(4) numerically, beginning at an altitude of 100 km, and calculating the decrease in altitude dh over a time interval dt according to $dh = -(v \sin \theta) dt$. The atmospheric density ρ_a is determined at each timestep by exponential interpolation from standard atmosphere tables³⁹. Once the central pressure $C_D \rho_a v^2/4$ exceeds the object's strength, equation (6) is used to calculate the changing radius r and effective cross-section $A = \pi r^2$.

Numerical results

Eight 'Tunguska bolides' are listed in Table 1. All cases correspond to an initial kinetic energy at $h = 100$ km of 15 Mton. (Objects lose different fractions of their kinetic energy before catastrophic disruption; for example, the 29-m stony asteroid entering at 45° has had its kinetic energy reduced by ablation and deceleration to ~ 10 Mton by the time it reaches 10 km altitude.) We compare five object types (iron, stony and carbonaceous asteroids; and short- and long-period comets), and, for stony asteroids, four incidence angles (15° , 30° , 45° and 90°). All other objects are incident at $\theta = 45^\circ$. The three asteroids have initial velocities of 15 km s^{-1} , the median impact velocity for Earth-crossing asteroids⁴⁰. Incident comet velocities are 25 km s^{-1} and 50 km s^{-1} for the short-period and long-period cases, respectively; again these are approximate median values^{40,41}. Both comets are assumed to have strengths of 10^6 dyn cm⁻², one-tenth that of the carbonaceous body.

Figure 1 compares energy release curves for the five kinds of object at $\theta = 45^\circ$. Figure 1a shows the altitude profile of energy release in units of Mton high explosive equivalent per kilometre. The two comets are entirely ablated, whereas the stony objects lose most of their kinetic energy to deceleration, not ablation. Figure 1b shows cumulative energy loss (Mton) along the object's path above the indicated altitude.

TABLE 1 Possibilities considered for the 15-Mton Tunguska object

Object type	Density (g cm ⁻³)	Mass (g)	Radius (m)	Velocity (km s ⁻¹)	Heat of ablation (erg g ⁻¹)	Yield strength (dyn cm ⁻²)	θ	Airburst height (km)
Iron	7.9	5.6×10^{11}	22	15	8×10^{10}	1×10^9	45°	0
Stone	3.5	5.6×10^{11}	29	15	8×10^{10}	1×10^8	45°	9
Carbonaceous	2.2	5.6×10^{11}	34	15	5×10^{10}	1×10^7	45°	14
SP comet	1.0	2.0×10^{11}	32	25	2.5×10^{10}	1×10^6	45°	23
LP comet	1.0	5.0×10^{10}	20	50	2.5×10^{10}	1×10^6	45°	29
Stone	3.5	5.6×10^{11}	29	15	8×10^{10}	1×10^8	15°	15
Stone	3.5	5.6×10^{11}	29	15	8×10^{10}	1×10^8	30°	11
Stone	3.5	5.6×10^{11}	29	15	8×10^{10}	1×10^8	90°	6

SP, short-period; LP, long-period.

Where the curves in Fig. 1a peak, objects have typically spread to ~ 5 –10 times their initial linear dimensions. This greatly enhances both ablation and deceleration, but it is misleading, because equation (2) ceases to be valid before an object has spread to this extent. Equation (2) assumes that the fragments of the disrupted object may continue to be treated with a single, collective bow shock. At some point the object spreads sufficiently for this approximation to break down, and the collective bow shock dissolves into separate bow shocks as the fragments accelerate away from one another³¹. This effect is presumably what produces crater-strewn fields³¹. It is not clear at what (spreading) radius greater than a bolide's initial radius the collective bow-shock approximation breaks down. This ambiguity does not greatly alter our conclusions regarding airburst altitude: once an object has spread to, say, twice its initial radius, its further spreading happens so quickly that an 'explosion altitude' is defined to within a few kilometres, regardless of whether the explosion is taken to occur then or when the object has spread to 5–10 times its initial radius.

The weak, fast-moving, easily ablated comets do not penetrate the atmosphere very deeply. Neither approaches the altitude of the Tunguska explosion. Even if the comets had strengths comparable to that of stony asteroids, they still could not fit the Tunguska observations. For example, giving the short-period comet a strength of $1 \times 10^8 \text{ dyn cm}^{-2}$ leads to complete ablation by an altitude of 16 km. Nor does assigning a 15-Mton incident comet an anomalously low asteroid-like velocity of 15 km s^{-1} allow it to penetrate below $\sim 16 \text{ km}$ before it is completely ablated. Moreover, comets may be less dense than the 1 g cm^{-3} value chosen here; values as low as $\rho_m \approx 0.3 \text{ g cm}^{-3}$ may be possible⁴². Such lower-density objects would airburst even higher.

The carbonaceous object, stronger and slower than the comets, penetrates more deeply, but also explodes at too high an altitude. If this carbonaceous bolide were to enter the atmosphere at $\theta = 60^\circ$, however, it would airburst at an altitude of 11–12 km.

These parameters are at the limits of those allowed by treefall simulations^{16,22}. Similarly, granting a carbonaceous object the strength of a stony asteroid would allow it to penetrate to comparable altitudes. Therefore a carbonaceous asteroid, although unlikely, cannot be ruled out as an explanation of the Tunguska event.

The stronger, denser, stony object in Table 1 penetrates to $\sim 9 \text{ km}$ above the ground, consistent with the Tunguska airburst altitude. Figure 2 shows the effect of impact angle on the explosion of stony asteroids. Results for four initial values of θ , 15° , 30° , 45° and 90° are given. The more nearly vertical an object's trajectory, the deeper it penetrates into the atmosphere before catastrophic disruption. Values of θ much less than 30° seem to be inconsistent with an explanation of the Tunguska object as a stony asteroid.

The most probable fate for a 15-Mton iron object is impact with the ground. Its high density (small cross-section) and high strength favour its survival. However, 15-Mton iron objects with speeds much in excess of the median value of 15 km s^{-1} will airburst before impact. Our simulations indicate that, although it is possible for such an iron object to reproduce the Tunguska event, an encounter speed in the range ~ 30 – 40 km s^{-1} is required. This excludes $\sim 90\%$ of known Earth-crossing asteroid orbits⁴⁰. These results are only weakly dependent on θ .

In summary, Tunguska was probably a fairly strong, dense, asteroid-like object, but probably not as strong or dense as iron. Carbonaceous asteroids and especially comets are unlikely candidates for the Tunguska object.

'Light nights' over Eurasia

Widespread 'light nights' were observed³ over Eurasia for the first few nights after the Tunguska event. Turco *et al.*¹² suggest three possible explanations: nightglow from NO_x – O_x reactions; dust; and noctilucent water clouds. The latter require injection or deposition at altitudes $> 50 \text{ km}$ to take advantage of the strong easterly winds needed to account for the observed geographical

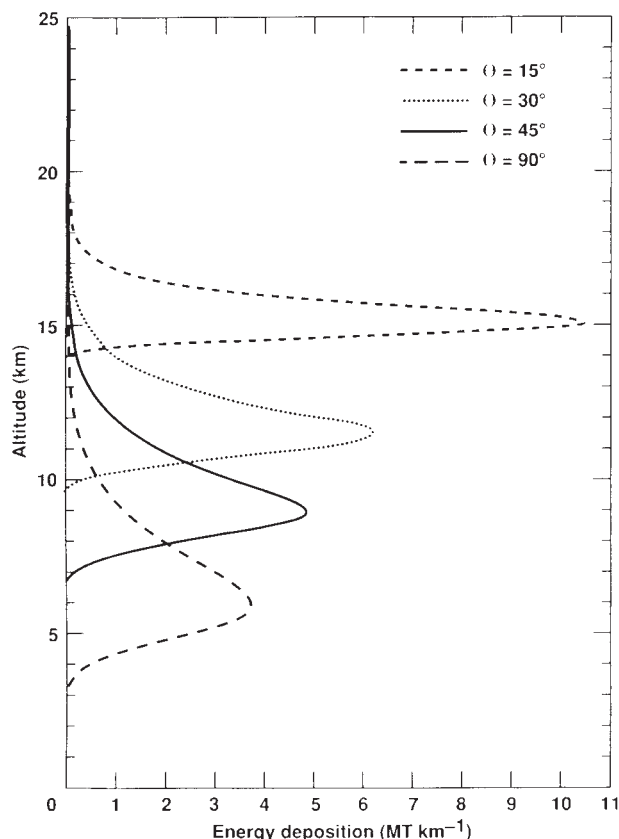


FIG. 2 Airburst altitudes for four 15-Mton stony asteroids with different incidence angles. The more nearly vertical an object's trajectory, the deeper it penetrates into the atmosphere before catastrophic disruption.

range of the phenomenon¹². The effect has also been attributed⁸ to the putative comet's tail, a less plausible explanation¹².

Tunguska was a ~15-Mton airburst at an altitude of roughly a scale height. The effects of explosion blast waves scale as the ratio of blast energy to atmospheric density⁴³. We therefore expect the Tunguska fireball to have reached a height above its airburst altitude appropriate to a $\sim 15e \approx 40$ Mton nuclear surface burst. The fireball for a 40-Mton surface burst rises to an altitude⁴³ of ~40 km. A rough estimate of the amount of water that could have been entrained by this fireball is $\pi R^2 H \rho_{\text{H}_2\text{O}}$, where R is the radius of the fireball when its pressure has dropped to the ambient value (several kilometres for a 10-Mton nuclear blast⁴⁴), and $\rho_{\text{H}_2\text{O}}$ the density of atmospheric water vapour (several times $10^{-3} \rho_a$). The result, $\sim 10^{35}$ molecules H_2O injected to ~50 km altitude, seems sufficient for noctilucent clouds to be produced¹². The Tunguska explosion may have lofted enough material high enough to account for the European 'light nights'.

Other impacts and recent airbursts

The model developed here for the Tunguska explosion may be tested for consistency with Meteor Crater in Arizona, and with the Tunguska-like explosions of the Revelstoke, Kincardine and other bolides. We consider these in turn.

Meteor Crater. Meteor Crater in Arizona was excavated $\sim 10^4$ yr ago by an iron object with a kinetic energy⁴ of ~15 Mton. Our model predicts that such an object disrupts so close to the surface that its fragments crater the surface as if from a coherent object, consistent with the appearance of Meteor Crater.

Revelstoke. In March 1965 a small bolide travelling at $\theta \approx 15^\circ$ exploded 30 km above Revelstoke, Canada¹³, with an energy in the tens of kiloton range⁴. Examination of recovered material shows it to have been a type I carbonaceous chondrite¹³. (Largely unprocessed ~1 mm sized fragments were recovered from the Revelstoke site¹³; only spherules of vaporized and recondensed material were recovered in the case of Tunguska³. Siderophile abundances in these spherules are consistent with extraterrestrial origin from a single object¹⁷, but are insufficient to determine the type of the Tunguska object. No scientific expeditions

reached the Tunguska site until almost 20 years after the event; it is not surprising that no other remnants of the object were found.) Microbarograph estimates of the explosive energy of the Revelstoke object range from 20 kton (ref. 4) to 70 kton (ref. 45).

We model Revelstoke as a carbonaceous chondrite (as in Table 1), moving at 15 km s^{-1} with an initial kinetic energy of 20 kton (corresponding to a cylinder of initial radius $r = 3.8 \text{ m}$) and $\theta = 15^\circ$. Our simulation shows that such an object's strength is exceeded at an altitude of ~33 km, giving maximum kinetic energy deposition to the atmosphere at ~32 km.

Kincardine and other bolides. The Kincardine and College objects exploded over Ontario in 1966 and Alaska in 1969, respectively, each at altitudes just over 60 km. Explosive energies in the kiloton to megaton range are possible⁴⁵. These objects therefore provide only weak constraints on our model. Our model does predict that long-period comets with kinetic energies in the tens of kiloton range should catastrophically disrupt and deposit the bulk of their energy just above 60 km altitude, so the fates of these objects are consistent with cometary airbursts.

The fates of bolides

The simple model for bolide deformation and catastrophic disruption presented here is consistent with the fates and explosive energies of the Meteor Crater, Tunguska and Revelstoke objects, provided that these objects were iron, stony and carbonaceous asteroids, respectively. The Meteor Crater and Revelstoke bolides are in fact known from recovered material to have had these identities. Within broad uncertainties, the explosions of the Kincardine and College objects are consistent with the fates of long-period comets with kinetic energies of tens of kilotons.

The fate of bolides in the Tunguska size range is strongly dependent on the nature of the object. Had the Tunguska object been a 15-Mton comet, far less surface destruction would have resulted, because of the much higher altitude of the airburst. Comets incident with tens of kilotons of energy explode so high in the atmosphere that they are scarcely noticed at the surface. Denser objects with higher material strengths are therefore necessarily more dangerous to surface life. □

Received 22 September; accepted 9 December 1992.

- Kulik, L. A. *Dokl. Acad. Nauk SSSR* **23**, 399–402 (1927).
- Baxter, J. & Atkins, T. *The Fire Came By* (Doubleday, Garden City NY, 1976).
- Whipple, F. J. W. *Q. J. R. met. Soc.* **56**, 287–304 (1930).
- Krinov, E. L. *Giant Meteorites* (Pergamon, Oxford, 1966).
- Shoemaker, E. M. A. *Rev. Earth planet. Sci.* **11**, 461–494 (1983).
- Cowan, C., Atluri, C. R. & Libby, W. F. *Nature* **206**, 861–865.
- Jackson, A. A. & Ryan, M. P. *Nature* **245**, 88–89 (1973).
- Fesenkov, V. G. *Sov. Astr.* **5**, 441–451 (1962).
- Petrov, G. I. & Stulov, V. P. *Cosmic Res.* **13**, 525–531 (1975).
- Liu, V. C. *Geophys. Res. Lett.* **5**, 309–312 (1978).
- Park, C. *Acta astronaut.* **5**, 523–542 (1978).
- Turco, R. P. *et al.* *Icarus* **50**, 1–52 (1982).
- Folinsbee, R. E., Douglas, J. A. V. & Maxwell, J. A. *Geochim. cosmochim. Acta* **31**, 1625–1635 (1967).
- Hunt, J. N., Palmer, R. & Penney, W. *Phil. Trans. R. Soc. Lond.* **252**, 275–315 (1960).
- Ben-Menahem, A. *Phys. Earth planet. Int.* **11**, 1–35 (1975).
- Zolotov, A. V. *Sov. Phys. Dokl.* **12**, 101–104 (1967).
- Ganapathy, R. *Science* **220**, 1158–1161 (1983).
- Rocchia, R. *et al.* in *Global Catastrophes in Earth History* (eds Sharpton, V. L. & Ward, P. D.), 189–193 (Geol. Soc. Am. SP-247, Boulder CO, 1990).
- Neftel, A., Beer, J., Oeschger, H., Zücher, F. & Finkel, R. C. *Nature* **314**, 611–613 (1985).
- Sekanina, Z. *Astr. J.* **88**, 1382–1414 (1983).
- Levin, B. Yu. & Bronshten, V. A. *Meteoritics* **21**, 199–215 (1986).
- Korobeinikov, V. P., Chushkin, P. I. & Shurshalov, L. V. *Acta astronaut.* **3**, 615–622 (1976).
- Zotkin, I. T. & Tsikulkin, M. A. *Sov. Phys. Dokl.* **11**, 183–186 (1966).
- Baldwin, B. & Sheaffer, Y. J. *Geophys. Res.* **76**, 4653–4668 (1971).
- Bronshten, V. A. *Physics of Meteoric Phenomena* (Reidel, Dordrecht, 1983).
- Melosh, H. J. *Impact Cratering* (Oxford Univ. Press, New York, 1989).

- Chyba, C. F., Thomas, P. J., Brookshaw, L. & Sagan, C. *Science* **249**, 366–373 (1990).
- Walker, J. C. G. *Icarus* **68**, 87–98 (1986).
- Zahnle, K. J. *J. geophys. Res.* **97**, 10,243–10,255 (1992).
- Biberman, L. M., Bronin, S. Ya. & Brykin, M. V. *Acta astronaut.* **7**, 53–65 (1980).
- Passey, Q. R. & Melosh, H. J. *Icarus* **42**, 211–233 (1980).
- Grigoryan, S. S. *Cosmic Res.* **17**, 724–740 (1979).
- Grieve, R. A. F. in *Geological Implications of Impacts of Large Asteroids and Comets on the Earth* (eds Silver, L. T. & Schultz, P. H.) 25–37 (Geol. Soc. Am. SP-190, Boulder CO, 1982).
- Seiff, A. in *Gas Dynamics in Space Exploration*, 19–32 (NASA SP-24, 1962).
- Öpik, E. J. *Physics of Meteor Flight in the Atmosphere* (Interscience, New York, 1958).
- Ivanov, B. A., Basilevsky, A. T., Kryuchkov, V. P. & Chernaya, I. M. *J. geophys. Res.* **91**, D413–D430 (1986).
- Sekanina, Z. in *Comets* (ed. Wilkening, L. L.), 251–287 (Univ. of Arizona, 1982).
- Hoerner, S. F. *Fluid-Dynamic Drag* (Hoerner, Midland Park NJ, 1965).
- NOAA *U.S. Standard Atmosphere*, 1976 (U.S. Govt Printing Office, Washington DC, 1976).
- Chyba, C. F. *Icarus* **92**, 217–233 (1991).
- Weissman, P. R. in *Global Catastrophes in Earth History* (eds Sharpton, V. L. & Ward, P. D.), 171–180 (Geol. Soc. Am. SP-247, Boulder CO, 1990).
- Peale, S. J. *Icarus* **82**, 36–49 (1989).
- Glasstone, S. & Dolan, P. J. *The Effects of Nuclear Weapons* (Dept of Defense, Washington DC 1977).
- Jones, E. M. & Kodis, J. W. in *Geological Implications of Impacts of Large Asteroids and Comets on the Earth* (eds Silver, L. T. & Schultz, P. H.) 175–186 (Geol. Soc. Am. SP-190, Boulder CO, 1982).
- ReVelle, D. O. *J. geophys. Res.* **81**, 1217–1230 (1976).

ACKNOWLEDGEMENTS. We thank D. Morrison, J. Rice, A. Seiff, D. Steel and O. B. Toon for discussions, and H. J. Melosh for a review. This work was supported in part by the National Research Council and NASA.

Extremely-high-velocity H₂O maser emission in the galaxy NGC4258

N. Nakai, M. Inoue & M. Miyoshi

Nobeyama Radio Observatory, Minamimaki, Minamisaku,
Nagano 384-13, Japan

WATER-VAPOUR maser emission at 22.23508 GHz has been observed in star-forming regions of our own galaxy and in the active nuclei of other galaxies¹⁻⁵. Maser emission from active galactic nuclei is typically 10⁵ times more luminous than in our galaxy, and the active galaxies concerned always exhibit radio or optical jets ejected from the nuclei. The superluminous masers are not well understood, but appear to be related to the mass outflow⁶. Using 16,000-channel spectrometers at the Nobeyama 45-m radio telescope, we have found H₂O maser emission in the nucleus of the galaxy NGC4258, offset by $\pm 1,000$ km s⁻¹ from the motion of the nucleus itself. This is much higher than the typical velocities of galactic masers (less than a few hundred kilometres per second). We suggest that the high-velocity maser emission in NGC4258 might be from masers orbiting a massive central black hole, or ejected in a bipolar outflow. More intriguingly, the features may not be Doppler in origin, but due to up- and down-shifted frequencies by Raman scattering in a dense plasma. Further high-resolution observations will be needed to distinguish between these possibilities.

NGC4258 is a spiral galaxy at a distance of 6.6 Mpc (ref. 7) with unusual 7-kpc-long bipolar jets or bubbles of radio continuum emission⁸⁻¹⁰. Extremely luminous water-vapour masers of ~ 100 times the luminosity of the Sun, $L_{\odot}$ (ref. 3), have been detected within 1 parsec (3×10^{18} cm) of its nucleus⁶. Our observations of the maser emission were made from January to June 1992 with the Nobeyama 45-m telescope. The detector was a cooled high-electron-mobility transistor receiver equipped with a circular polarizer, combined with 2,048-channel acousto-optical spectrometers (AOS). We used an array of eight such AOSs arranged to detect different frequencies, resulting in a total frequency range of 285 MHz and a velocity coverage of 3,850 km s⁻¹. The spatial resolution of the antenna was 74".

Figure 1 shows the spectrum of H₂O maser emission towards the centre of NGC4258. In addition to the known central features at the systemic velocity of the galaxy (corresponding to the velocity of its centre $v_{\text{sys}} = 460$ km s⁻¹), there are strong redshifted features with intensities of about one-half or one-third

of the main features at $v_{\text{LSR}} = 1,200$ – $1,440$ km s⁻¹ (where LSR stands for local standard of rest) and weak blueshifted features with intensities of about one-tenth at -300 to -460 km s⁻¹. The velocities relative to the systemic velocity ($v_{\text{rel}} \equiv v_{\text{LSR}} - v_{\text{sys}}$) are 740 to 980 km s⁻¹ and -760 to -920 km s⁻¹, respectively. In the spectrum obtained in June 1992, another component at $v_{\text{LSR}} = -525$ km s⁻¹ ($v_{\text{rel}} = -985$ km s⁻¹) was detected, with a flux density of 59 mJy. No significant emission stronger than 51 mJy (3σ) could be recognized outside the velocity range in Fig. 1. The linewidth of individual components of the high-velocity features is $\Delta v_{\text{FWHM}} = 1$ – 6 km s⁻¹. The features varied in intensity during the observing period. In the observed frequency range, no other molecular or recombination line has been detected so far in interstellar space. Molecular lines of HDO and HO³⁷Cl are expected at frequencies corresponding to $v_{\text{rel}} = -971$ and 912 km s⁻¹, respectively, but these lines have not yet been detected in space and must be very weak. These facts indicate that the extremely-high-velocity features are actually H₂O maser emission from NGC4258.

Spatial mapping of the H₂O maser emission showed that the redshifted features lie within 12" (380 pc) of the nucleus of the galaxy. The blueshifted features were too weak for us to measure their positions. Furthermore, we have made very-long-baseline interferometer (VLBI) observations of the main features and of the strongest redshifted feature at $v_{\text{LSR}} = 1,322$ km s⁻¹, by using the 45-m and the Kashima 34-m telescopes. The result indicates that the 1,322 km s⁻¹ feature is located within $\pm 0.05''$ (± 1.6 pc) from the main features.

The isotropic luminosity of individual high-velocity components is calculated to be $L = 0.1$ – $6 L_{\odot}$, which is comparable to the total luminosity of the most powerful maser source in our Galaxy, W49N, and the total luminosities of redshifted and blueshifted features are $23 L_{\odot}$ and $1 L_{\odot}$, respectively. The total luminosity of the main feature is $101 L_{\odot}$.

Figure 2 shows the time variation of the flux densities and velocities at peak intensities for the strongest four redshifted components at $v_{\text{LSR}} = 1,290$, 1,322, 1,385 and 1,435 km s⁻¹. The synchronized variations between the redshifted components in the flux and the peak velocity suggest that these maser components have a common pumping source and are distributed within ~ 0.1 pc each other. The variability of the shifted features does not seem to correlate with that of the main features. The linewidths of the components of 1,290, 1,322, and 1,435 km s⁻¹ remained nearly constant for half a year, indicating saturated maser emission, but the line-width of the 1,385 km s⁻¹ component varied with the flux density, S , roughly as $S^{-0.5}$,

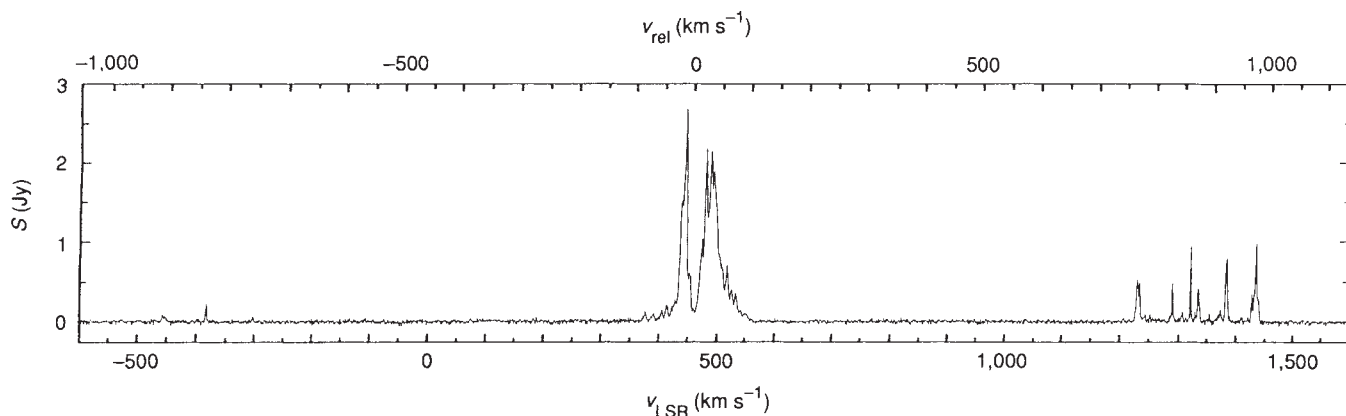


FIG. 1 Spectrum of H₂O maser emission towards the centre of NGC4258 (right ascension $\alpha = 12^{\text{h}} 16^{\text{m}} 29.36^{\text{s}}$, declination $\delta = 47^{\circ} 34' 53''$, in 1950.0)⁶ displayed from $v_{\text{LSR}} = -600$ to $1,600$ km s⁻¹, although the observed velocity range is $v_{\text{LSR}} = -1,200$ to $+2,650$ km s⁻¹. No significant

emission can be recognized outside the velocity range displayed. The upper scale is the velocity relative to the centre of NGC4258 (moving at 460 km s⁻¹).

suggesting an unsaturated maser. The variability of the blue-shifted features could not be measured, because their intensity was too low.

Velocities up to $\pm 1,000 \text{ km s}^{-1}$ are well out of the range of the rotation velocity ($\pm 250 \text{ km s}^{-1}$) of the galaxy NGC4258⁷ and are much higher than any velocities known for molecular gas in our Galaxy or other galaxies (for example $\parallel v \parallel = 200 \text{ km s}^{-1}$ for H_2O maser in W49N¹¹; $200\text{--}300 \text{ km s}^{-1}$ for CO in the galaxy M82¹²; 170 km s^{-1} for OH in the galaxy NGC253¹³). The true velocities may be higher than the velocities measured in projection. To explain such extreme velocities, we consider three possibilities.

The first is that the masers are emitted from a circumnuclear molecular torus rapidly rotating around a compact massive object, such as a black hole. Assuming a radius of 1 pc for the torus, a mass of $2 \times 10^8 M_\odot$ for the nuclear object can produce keplerian motion at $1,000 \text{ km s}^{-1}$. In this case, the explanation for the stronger main features, for which $|v_{\text{rel}}| \leq 150 \text{ km s}^{-1}$, must match that for the weaker high-velocity features.

The second possibility is that the high-velocity masers are emitted in molecular gas flowing out from (or flowing in towards) the nuclear region. If this is a bipolar flow perpendicular to the galactic disk, the outflow velocity becomes $\sim 3,000 \text{ km s}^{-1}$, corrected for the inclination angle of the disk (72°)¹⁴. Such a mass outflow from the nuclear region of NGC4258 cannot be ruled out, because strong radio continuum emission and optical jets or bubbles are emitted from the nucleus^{8–10,15–17}, indicating some outflow mechanism at work in the region. Recently optical studies¹⁸ of the galaxy reveal velocities as high as 600 km s^{-1} , which have been interpreted as ballistic motions of $2,000 \text{ km s}^{-1}$ (after projection effects are taken into account). The bond strength of O–H in a water molecule is $119 \text{ kcal mol}^{-1}$ or 5.2 eV , corresponding to the kinetic energy of the molecule if it moves at only 7.4 km s^{-1} . Thus for the molecule to survive rapid motion at $\sim 1,000 \text{ km s}^{-1}$ or $\sim 3,000 \text{ km s}^{-1}$, the molecular gas must be gradually accelerated, so that the conversion of the energy to the internal motion of molecule by external shocks is very small. The main maser features may have originated from the slowly rotating disk or from the outflow gas near the galactic nucleus.

The third possibility is that the shifted features are not due to the Doppler motion of the gas but due to stimulated Raman scattering of the main features. Stimulated Raman scattering in a plasma has been investigated in terms of forward scattering¹⁹, shock waves²⁰ and (recently) backward scattering²¹, and the results have been applied to high-velocity features of water masers. If the maser radiation of the main features intersects dense, plasma-like, compact ionized regions, the stimulated Raman scattering in the plasma (due to nonlinear interaction between the longitudinal and transverse waves) is expected to produce additional waves with up- and down-shifted frequencies, which are separated by the plasma frequency from the original one. Although the shifted maser could also produce emission at a frequency separated from the main features by twice the plasma frequency, no significant emission stronger than 51 mJy (3σ) is found at $v_{\text{LSR}} = 1,940\text{--}2,420 \text{ km s}^{-1}$ [$v_{\text{rel}} = 2 \times (740\text{--}980 \text{ km s}^{-1})$]. Deguchi²¹ has made a detailed quantitative analysis of stimulated Raman backward scattering for the astrophysical conditions. According to his calculations, it is possible to create the Raman instability in plasma with an electron density of $\sim 10^7 \text{ cm}^{-3}$ (the value for the extragalactic masers), overwhelming the loss of photons due to free-free absorption.

The bidirectional collimated ejection in the nuclear region of NGC4258, indicated by radio continuum and optical jets, favours the second possibility, whereas the symmetry of the shifted maser features with respect to the central velocity favours the model of stimulated Raman scattering. The bipolar outflow model does however raise the question of how molecules can survive in a strong shock in a medium with the extremely high velocity of $\geq 1,000 \text{ km s}^{-1}$ (for comparison, the sound velocity in interstellar space is only a few kilometres per second). The stimulated Raman scattering, on the other hand, requires some special conditions. For example, the efficiency of the production of scattering photons must be unusually high, $0.01\text{--}0.1$, in the dense plasma. To produce the blueshifted (higher-frequency) features with observed fluxes, another maser beam must interact with the main maser at an intersecting angle of nearly 180° (ref. 21). The models must account for the absence of significant maser emission ($< 51 \text{ mJy} = 3\sigma$) between the main and shifted

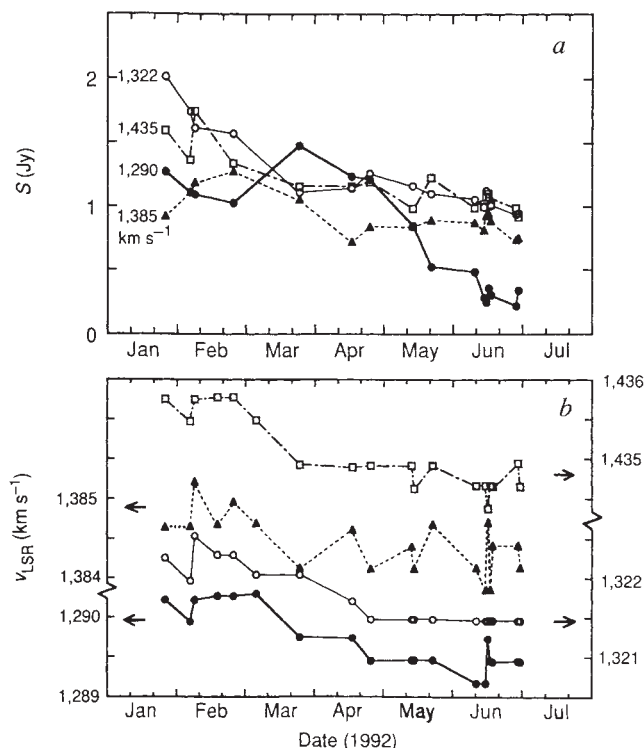


FIG. 2 Time variation of the flux densities (a) and the velocities at peak intensities (b) for the strongest four redshifted components at $v_{\text{LSR}} \approx 1,290$, $1,322$, $1,385$ and $1,435 \text{ km s}^{-1}$. Although the relative strengths of these components varied, the absolute flux densities of all components generally decreased during January to June. This trend does not seem to correlate with the variability of the main features. The peak velocities of all four components also shifted by $\sim 1 \text{ km s}^{-1}$ from January to April and were almost constant from April to June.

features. Further observational (especially VLBI) and theoretical investigations are required to discriminate between models of the extremely-high-velocity features. Results for other galaxies, such as NGC3079, will be reported elsewhere. $\square$

Received 7 September; accepted 17 November 1992.

1. dos Santos, P. M. & Lepine, J. R. D. *Nature* **278**, 34–35 (1979).
2. Gardner, F. F. & Whiteoak, J. B. *Mon. Not. R. astr. Soc.* **201**, 13p–15p (1982).
3. Claussen, M. J., Heiligman, G. M. & Lo, K. Y. *Nature* **310**, 298–300 (1984).
4. Henkel, C. *et al. Astr. Astrophys.* **141**, L1–L3 (1984).
5. Haschick, A. D. & Baan, W. A. *Nature* **314**, 144–146 (1985).
6. Claussen, M. J. & Lo, K.-Y. *Astrophys. J.* **308**, 592–599 (1986).
7. Burbidge, E. M., Burbidge, G. R. & Prendergast, K. H. *Astrophys. J.* **138**, 375–384 (1963).

8. van der Kruit, P. C., Oort, J. H. & Mathewson, D. S. *Astr. Astrophys.* **21**, 169–184 (1972).
9. de Bruyn, A. G. *Astr. Astrophys.* **58**, 221–236 (1977).
10. van Albada, G. D. & van der Hulst, J. M. *Astr. Astrophys.* **115**, 263–269 (1982).
11. Morris, M. *Astrophys. J.* **210**, 100–107 (1976).
12. Nakai, N. *et al. Publ. astr. Soc. Japan* **39**, 685–708 (1987).
13. Turner, B. E. *Astrophys. J.* **299**, 312–333 (1985).
14. van Albada, G. D. *Astr. Astrophys.* **90**, 123–133 (1980).
15. Courtes, G. & Gruveller, P. C. R. *Acad. Sci. Paris* **253**, 218 (1961).
16. Deharveng, J. M. & Pellet, A. *Astr. Astrophys.* **9**, 181–188 (1970).
17. van der Kruit, P. C. *Astrophys. J.* **192**, 1–19 (1974).
18. Cecil, G., Wilson, A. S. & Tully, R. B. *Astrophys. J.* **390**, 365–377 (1992).
19. Fernandez, J. C. & Reinisch, G. *Astr. Astrophys.* **67**, 163–174 (1978).
20. Burdjuza, V. V., Charugin, V. M. & Tomozov, V. M. *Astr. Astrophys.* **79**, 306–311 (1979).
21. Deguchi, S. *Astrophys. J.* (submitted).

ACKNOWLEDGEMENTS. We thank S. Deguchi and K. Kawaguchi for discussions, K. Nishiyama for help with observations with the 45-m telescope, and H. Takaba and T. Iwata for the VLBI observations.

A long-period globular-cluster pulsar in an eclipsing binary system

A. G. Lyne, J. D. Biggs, P. A. Harrison & M. Bailes

University of Manchester, Nuffield Radio Astronomy Laboratories,
Jodrell Bank, Macclesfield, Cheshire SK11 9DL, UK

PULSARS that are members of binary systems in globular clusters are all rapidly rotating, and it is assumed that they have been spun up by accretion from binary companions. Here we report an exception: PSR1718–19, in the globular cluster NGC6342, is in a 6.2-hour eclipsing binary system, but has the relatively long period of 1 s. Its magnetic field is strong, $\sim 10^{12}$ G, and its spin-down age is small, ~ 10 Myr. Furthermore, the mass of its companion is only 0.1–0.2 solar masses ($M_{\odot}$). The eclipses show that the binary system is embedded in a cloud of material which must have been ejected from the companion star, although calculations suggest that the companion is well inside its Roche lobe. The pulsar's radiation may be causing expulsion of material beyond the Roche lobe, as in the ablating binary systems containing PSR-1744–24A and PSR1957+20, even though the incident flux at the companion is orders of magnitude smaller than in these cases. Pulsars may therefore have a much greater influence on their companions than has been supposed.

The binary radio pulsar PSR1718–19 was discovered in 1990 in a survey of globular clusters at Jodrell Bank^{1,2} at a frequency of 610 MHz. Subsequent measurements were carried out with the 76-m Lovell telescope at frequencies between 408 MHz and 1,660 MHz. Arrival times were obtained by convolution of the pulses with an appropriate template and then corrected for the Earth's motion using the JPL DE200 barycentric ephemeris³. A simple model for the pulsar period, period-derivative, position and binary motion was fitted to the barycentric arrival times⁴ (Table 1).

The pulsar is in a very short period, circular orbit. Although the masses m of the two objects cannot be determined uniquely, some information can be obtained from the mass function, calculated from only two observables, the projected pulsar semi-major axis $a_p \sin i$, and the orbital period, P_b :

$$f(m_p, m_c) = \frac{4\pi^2 (a_p \sin i)^3}{G P_b^2} = \frac{(m_c \sin i)^3}{(m_p + m_c)^2} \\ = 0.000706 M_{\odot} \quad (1)$$

where subscripts p and c refer to the pulsar and companion respectively and i is the unknown inclination angle of the orbit to the plane of the sky. Although the inclination angle cannot be determined, the *a priori* probability that it is less than i_0 is $1 - \cos i_0$. Therefore, if we assume that the pulsar has the canonical neutron star mass of $m_p = 1.4 M_{\odot}$, m_c must be at least 0.12 $M_{\odot}$ and has 95% probability of being less than 0.43 $M_{\odot}$. The corresponding separation of the objects, $a = a_p + a_c$, is about 1.4×10^6 km and is almost independent of inclination.

During the course of the confirming observations, it became clear that the flux density of the pulsar was varying throughout the orbit, apparently because of partial eclipse by material associated with the companion star. The form of the eclipses was determined at 408, 606, 1,404 and 1,660 MHz. The results, averaged over several orbits at each frequency, are summarized in Fig. 1. Although the eclipse is only partial at the two highest frequencies, so that the pulses are observed throughout the orbit, strong pulses are seen for only about one-quarter of the time at 606 MHz, close to phase 0.75 when the pulsar is nearest to the Earth. There was only a weak detection at 408 MHz, again at about the same phase. The lack of any hard eclipse at high frequencies suggests that the eclipse is due to scattering or absorption of the radio waves as they pass through ionized matter surrounding the companion star. Although the spectral index at phase 0.7, close to the maximum, is about -2.0 between 606 MHz and 1,660 MHz and typical of most pulsars, at 408 MHz the flux density is actually less than that at 606 MHz

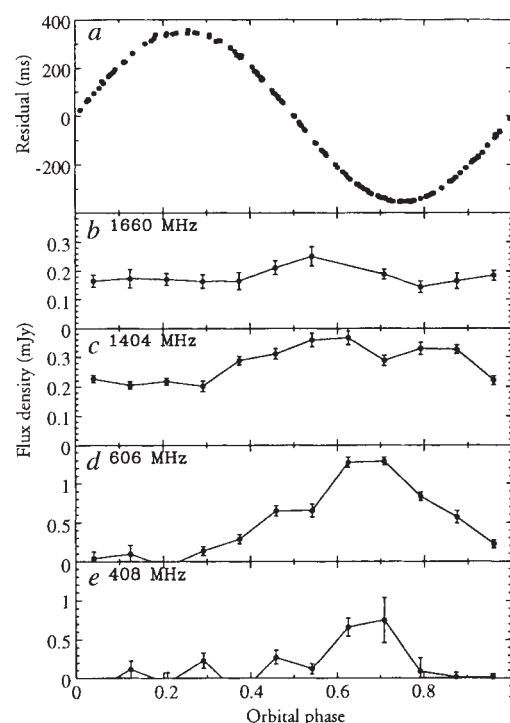


FIG. 1 a, Timing residuals for PSR1718–19, obtained assuming the simple slowdown model given in Table 1, plotted as a function of orbital phase. The advance and retardation of the pulses can be seen as the pulsar orbits the companion. The pulsar is closest to the Earth at phase 0.75 and behind the companion at phase 0.25. b–e, Time-averaged pulse flux density as a function of orbital phase at frequencies of 1,660, 1,404, 606 and 408 MHz respectively.

TABLE 1 Measured and derived parameters of the PSR1718–19 system

Epoch (MJD)	48,147.5
Pulsar period, P	$1.004037129 \pm 0.000000001$ s
Pulsar period derivative, $\dot{P}$	$(1.59 \pm 0.02) \times 10^{-15}$
Dispersion measure, DM	71.0 ± 0.2 cm $^{-3}$ pc
Right ascension (1950.0)	17 hr 18 min 04.76 s ± 0.16 s
Declination (1950.0)	$-19^{\circ}33'18'' \pm 18''$
Projected semi-major axis, $a_p \sin i$	352.6 ± 0.8 light ms
Orbital period, P_b	$22,314.83 \pm 0.02$ s
Eccentricity, e	≤ 0.005
Assumed longitude of periastron, ω	0.0°
Time of ascending node (MJD), T_0	48455.0237 ± 0.0001 days
Peak flux density at 408 MHz	0.7 mJy
Peak flux density at 606 MHz	1.3 mJy
Peak flux density at 1404 MHz	0.36 mJy
Peak flux density at 1,660 MHz	0.26 mJy
Pulsar characteristic age, $\tau_c = P/2\dot{P}$	1.0×10^7 yr
Pulsar surface magnetic flux, B	1.5×10^{12} G
Mass function, $f(m_p, m_c)$	$7.06 \times 10^{-4} M_{\odot}$
Companion mass, m_c	$\geq 0.12 M_{\odot}$
Pulsar-companion separation, a	$\sim 1.4 \times 10^6$ km

and is lower than the extrapolated spectrum by a factor of about 4 (Table 1), indicating that strong absorption or scattering is occurring even when the companion is on the far side of the pulsar. This and the long duration of the eclipse at low frequency require that this material extends well beyond the pulsar and of course the companion's Roche lobe, the region within which matter is gravitationally bound to the star.

Although Roche-lobe overflow from the companion star seems the most likely source of this material, there is a problem. The radius R_L of the Roche lobe is given by⁵

$$R_L = \frac{0.49a}{0.6 + (m_p/m_c)^{2/3} \ln(1 + (m_p/m_c)^{-1/3})} \quad (2)$$

If the companion is a main-sequence star, then its radius R_c is approximately given by⁶ $R_c/R_{\odot} = m_c/M_{\odot}$. The eclipses show that $i \geq 45^{\circ}$, in which case the ratio of R_c to R_L is ≥ 0.4 , so that the surface of the companion is well inside its Roche lobe and there should be no significant mass loss. If the companion is a helium main-sequence or white dwarf star, it will have an even smaller radius for a given mass and the conclusion remains valid.

One possibility is that the eclipsing material results from the ablation of the surface of the companion star by radiation from the pulsar, derived from its spin-down energy⁷. This has been proposed to explain the eclipses seen in the two other eclipsing pulsars with low-mass companions, PSR1957+20 (ref. 8) and PSR1744-24A (ref. 9).

In the case of PSR1718–19, however, the pulsar has comparatively little kinetic energy, and in particular, the present pulsar rotational energy loss rate of $\dot{E} = I\omega\dot{\omega}$ provides a maximum energy flux density at the companion star of only 3×10^8 erg s $^{-1}$ cm $^{-2}$, assuming a moment of inertia $I = 10^{45}$ g cm 2 for the neutron star and that the energy is radiated isotropically. This is many orders of magnitude less than that of PSR1957+20 and probably about two orders of magnitude less than our best estimate for that of PSR1744–24A, although $\dot{P}$ is unknown because of gravitational contamination by the cluster. The energy flux of PSR1718–19 at the surface of the companion is a few orders of magnitude below the critical flux $F_{Ch} = 10^{11} - 10^{12}$ erg cm $^{-2}$ necessary for substantial swelling of the companion due to irradiation¹⁰. This flux may be enhanced by beaming effects, but would still be insufficient to achieve F_{Ch} . It does seem, however, that Roche-lobe overflow is taking place, and that pulsar radiation is responsible and may play a much larger role in determining the evolution of companion stars than previously thought¹⁰.

There is substantial evidence that the pulsar is associated with the globular cluster NGC6342. The pulsar lies within about 2.3 arcmin of the centre of the target cluster, well within the tidal radius of 9 arcmin, although it is outside the core radius¹¹ of 8 arcsec. The cluster is believed¹¹ to be 11.6 kpc away, at galactic latitude $b = +9.7^{\circ}$, and its distance above the galactic plane is about 2 kpc. Observations of other pulsars in globular clusters show that the layer of ionized gas causing dispersion extends to only about 1 kpc (ref. 12) and gives a total dispersion measure contribution of $DM \approx 20/\sin b$ cm $^{-3}$ pc. In this case, the expected DM is therefore ~ 120 cm $^{-3}$ pc, a little high but reasonably consistent with the observed value of 71 cm $^{-3}$ pc, given the large spread of $DM \sin b$ in other clusters. Furthermore, the proportion of pulsars with binary companions is two orders of magnitude greater in globular clusters than in the galactic disk. In conclusion, the position, distance and binary nature all suggest that the system is probably associated with the cluster.

The measured period derivative of PSR1718–19 gives a characteristic age of ~ 10 Myr. In the absence of any accretion, this is generally accepted as being an upper limit to the age of the pulsar⁴, indicating that the pulsar is young compared with other globular cluster pulsars. The large value of $\dot{P}$ also indicates that the pulsar has a surface magnetic flux density of 1.3×10^{12} G, about two orders of magnitude greater than any other cluster pulsar and more typical of pulsars in the galactic plane. At its distance from the cluster core, any contribution to the period derivative from acceleration in the cluster gravitational field is negligible¹³.

What is the origin of this unusual system? Because star formation and hence the formation of neutron stars by supernova collapse in NGC6342 ceased long ago, the apparent youth of the pulsar suggests that a major event has occurred at some time during the last 10 Myr. The large 1-s spin period is close to its equilibrium value⁶, suggesting Eddington-limited accretion at some time in the near past. The most likely formation path therefore seems to be collision of an old neutron star with a main-sequence star ~ 10 Myr ago, leading to catastrophic mass transfer to the neutron star and considerable loss of matter from the system. An alternative is that the neutron star was in fact formed recently, in the accretion-induced collapse of a white dwarf following accretion from a lower-mass companion^{6,14}. Such a collapse is accompanied by sudden mass-loss, giving the system a velocity impulse which could result in ejection from the cluster core. Any orbital eccentricity produced by the collapse must have been subsequently removed by tidal forces.

If PSR1718–19 is not located in the globular cluster, then the neutron star is probably young and the system's origin could be similar to that of the Her X–1 system. In this scenario (ref. 6 and references therein), the system evolved from a close binary system consisting of a $\sim 15 M_{\odot}$ star and a $\sim 2 M_{\odot}$ star. The more massive star evolved and expanded in its giant phase, resulting in a common-envelope phase that reduced the orbital separation, and then underwent a supernova explosion without disrupting the system. Finally, in the PSR1718–19 system, the orbit was circularized and reduced to such an extent that the pulsar's radiation could cause substantial mass loss from the companion.

If the pulsar is a rejuvenated old neutron star, the large magnetic field suggests that there has been little or no decay from its original value, lending support to work suggesting the field decay may not occur^{15–17}, in contrast to other evidence^{18–20}. On the other hand, if it is a galactic pulsar or if it were formed by accretion-induced collapse, it may be young and the large magnetic field would not be as surprising. □

Received 21 October 1992; accepted 17 November 1992.

- Biggs, J. D., Lyne, A. G. & Brinklow, A. in *Timing Neutron Stars*, (Nato ASI Series) (eds Ögelman, H. & van den Heuvel, E. P. J.) 157–162 (Kluwer, Dordrecht, 1989).
- Biggs, J. D., Lyne, A. G. & Johnston, S. in *Proc. 23rd ESLAB Symp. Two Topics in X-ray Astronomy* (eds Hunt, J. & Batrick, B.) 293–299 (ESA, Noordwijk, 1989).
- Standish, E. M. *Astr. Astrophys.* **114**, 297–302 (1982).

4. Manchester, R. N. & Taylor, J. H. *Pulsars* (Freeman, San Francisco, 1977).
5. Eggleton, P. P. *Astrophys. J.* **268**, 368 (1983).
6. Bhattacharya, D. & van den Heuvel, E. P. J. *Phys. Rep.* **203**, 1-124 (1991).
7. Ruderman, M., Shaham, J. & Tavani, M. *Astrophys. J.* **336**, 507-518 (1989).
8. Fruchter, A. S., Stinebring, D. R. & Taylor, J. H. *Nature* **333**, 237-239 (1988).
9. Lyne, A. G. *et al. Nature* **347**, 650-652 (1990).
10. Podsiadlowski, P. *Nature* **350**, 136 (1991).
11. Webbink, R. F. in *Dynamics of Star Clusters*, IAU Symp. No 113 (eds Goodman, J. & Hut, P.) 541-577 (Reidel, Dordrecht, 1985).
12. Bhattacharya, D. & Verbunt, F. *Astr. Astrophys.* **242**, 128-132 (1991).
13. Phinney, E. S. *Mon. Not. R. astr. Soc.* (in the press).
14. Michel, F. C. *Nature* **329**, 310-312 (1987).
15. Taam, R. E. & van den Heuvel, E. P. J. *Astrophys. J.* **305**, 235-245 (1986).
16. Bailes, M. *Astrophys. J.* **342**, 917-927 (1989).
17. Bhattacharya, D., Wijers, R. A. M. J., Hartman, J. W. & Verbunt, F. *Astr. Astrophys.* **245**, 198-212 (1992).
18. Lyne, A. G., Anderson, B. & Salter, M. J. *Mon. Not. R. astr. Soc.* **201**, 503-520 (1982).
19. Lyne, A. G., Manchester, R. N. & Taylor, J. H. *Mon. Not. R. astr. Soc.* **213**, 613-639 (1985).
20. Harrison, P. A., Lyne, A. G. & Anderson, B. *Mon. Not. R. astr. Soc.* (in the press).

ACKNOWLEDGEMENTS. We thank C. Robinson for help with data analysis.

Evidence for heterogeneous reactions in the Antarctic autumn stratosphere

J. G. Keys*, P. V. Johnston*, R. D. Blatherwick† & F. J. Murcray†

* National Institute of Water and Atmospheric Research, Lauder, Central Otago, New Zealand

† Physics Department, University of Denver, Colorado 80208, USA

REACTIVE chlorine compounds are known to cause ozone depletion in the Antarctic stratosphere, but they can be bound into an inactive form through reactions with nitrogen dioxide. In the spring, NO₂ can be converted to a long-lived reservoir species, HNO₃, on the surface of polar stratospheric clouds^{1,2}. This removes NO₂ from the stratosphere and allows chlorine-catalysed ozone destruction to proceed. It has been suggested that similar reactions may take place on background sulphate aerosols in the Antarctic stratosphere³, but as yet there has been no unambiguous evidence for these reactions in the absence of polar stratospheric clouds (although there have been observations of ozone loss attributed to volcanic aerosols^{4,5}). Here we present measurements of Antarctic stratospheric NO₂ and HNO₃ concentrations taken in 1991. Our results demonstrate that reactive nitrogen was converted to HNO₃ in autumn, before temperatures were low enough for polar stratospheric clouds to form. We conclude that heterogeneous chemistry on background aerosols was responsible for this conversion, which brought with it the potential for additional ozone loss in the autumn.

Antarctic ozone depletion is believed to be dominated by reactive chlorine in the stratosphere participating in a catalytic cycle involving the ClO dimer, Cl₂O₂ (ref. 6). The importance of NO₂ in the Antarctic stratosphere stems from its influence on the partitioning of chlorine between the reactive form, ClO, and the less reactive form, chlorine nitrate, ClONO₂ (see for example ref. 7):



Here we examine seasonal changes in NO₂, and in its long-term storage reservoir, HNO₃. Figure 1 shows the history of vertical column NO₂ during 1991. The data were acquired using ground-based differential absorption spectroscopy at Arrival Heights, Antarctica (78° S, 167° E). In autumn, there is a progressive loss of NO₂ as the nights grow longer. When the Sun first reaches solar zenith angles greater than 95° near day 55, continuous daytime photolysis of NO₃ ceases, allowing the production of N₂O₅ through the reactions



NO₂ column amounts remain low during winter and in the springtime polar vortex, and later in spring gradually recover to their summer maximum values. Superimposed on the general slope of the springtime recovery are large cyclic increases in NO₂, which are a consequence of planetary wave activity moving the NO₂ and ozone depleted vortex away from the observing site.

In January 1991 a Fourier transform infrared spectrometer (FTIR) was installed at the Arrival Heights laboratory to measure HNO₃ in absorption, and in particular to observe the behaviour of the nitric acid column in autumn as the Sun left the polar cap. Measurements were made throughout the year as conditions permitted. Data were analysed by comparing observed and synthetic spectra for the ν₃ bands of HNO₃ in the 866 cm⁻¹ to 870 cm⁻¹ spectral region, where there are seven well defined manifolds of HNO₃ and an absence of other absorbers. Synthetic spectra were generated using the latest line strengths available (A. Goldman, personal communication) and ray tracing through a multilayer model of the atmosphere.

Results of the autumn and spring nitric acid measurements are shown in Fig. 2, together with total column ozone data. It can be seen that during autumn there is a steady increase in the HNO₃ column amount, followed by a decline over winter and a recovery in late spring. The conventional understanding of nitric acid photochemistry is that HNO₃ is produced by the reaction of NO₂ with OH, and destroyed by photolysis and by reaction with OH. In the Antarctic stratosphere, the lifetime is likely to be many weeks. As autumn progresses, the gas-phase model predicts no substantial increase of HNO₃ (refs. 8, 9). A simulated conversion of N₂O₅ to HNO₃ on background aerosols, however, shows column nitric acid to increase between days 55 and 75 in autumn by about 3.0 × 10¹⁵ molecule cm⁻² (ref. 9). Those calculations also show that the overnight decay of NO₂ at Arrival Heights agrees best with the model when the effect of background aerosols is included. A straight-line fit to the nitric acid data of Fig. 2 present between day 55 (when N₂O₅ formation starts) and the end of the season shows a rate of increase in the column amount that is very close to the model calculations above. The NO₂ and HNO₃ data of Figs 1 and 2 are therefore consistent with the view that the autumn increase in HNO₃ is due to a surface-catalysed reaction on background aerosols¹⁰



Laboratory work has shown that the N₂O₅ uptake in this reaction is surprisingly fast, and fairly insensitive to the sulphuric acid concentration in the aerosol or the temperature¹¹. Perhaps more important for data interpretation, the nitric acid is believed to

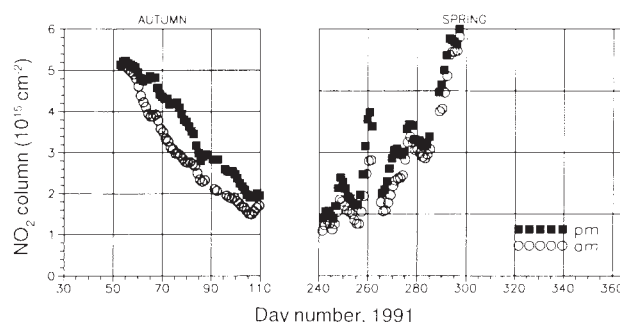


FIG. 1 Three-day running averages of NO₂ vertical column amounts for 1991 in units of molecules cm⁻², for morning and evening when the Sun is at a solar zenith angle of 90°. The lower morning values are a consequence of overnight conversion of NO₂ to other reservoirs of NO_x. The seasonal trend in this overnight decay of NO₂ is best fitted by model results when the effects of reactions on background aerosols are included⁹. The approximate vertical column has been obtained by dividing the slant column by 20.

be returned to the gas phase¹², and is therefore measurable using ground-based FTIR techniques, unlike heterogeneous reactions on polar stratospheric cloud surfaces in the Antarctic spring where HNO_3 can be sequestered in the solid phase. By the end of autumn, then, most of the NO_x can be expected to be in the form of nitric acid. By the beginning of spring, the column amount of HNO_3 has fallen to $\sim 1.6 \times 10^{16}$ molecule cm^{-2} , about half the value at the end of autumn. The temperature at 20 km altitude on day 240 was 186 K (ref. 13). Our interpretation is that gas-phase HNO_3 has by that time been sequestered in the condensed phase by heterogeneous reactions on polar stratospheric cloud surfaces, forming nitric acid trihydrate as temperatures fell below 196 K in winter



It is probable that a substantial fraction of nitric acid has not only been lost from the gas phase, but also removed from the stratosphere by the growth and subsequent sedimentation of aerosol particles. Evidence of such 'denitrification' is provided in Antarctic measurements from aircraft^{14,15}. Figure 2 shows that for spring there was a cyclical pattern of peaks and troughs in both O_3 and HNO_3 , with peaks for example near days 260 and 275. The peaks also correlate with high NO_2 values, so our assumption is that in these episodes we are measuring lower-latitude air as the denitrified polar vortex moves away from the site under the influence of planetary wave activity, as has been demonstrated with previous NO_2 data¹⁶. The nitric acid amount as measured inside the vortex reaches its lowest levels of $\sim 1.1 \times 10^{16}$ molecule cm^{-2} near day 287, and it can be seen in Fig. 2 that this coincides with minimum ozone amounts of about 140 Dobson units (DU). The HNO_3 column amount measured inside the vortex is in close agreement with values observed in 1986 (ref. 17), and a little higher (within the errors) than that recorded from the Arrival Heights laboratory in 1987. Considerable variability can be expected during spring and from year to year, both because the location of the vortex varies with respect to the observing site, and because of the large horizontal gradient in HNO_3 across the vortex edge^{15,19}.

As spring progresses and the days lengthen, the O_3 and NO_2 column amounts recover, and at the same time there is a steady increase in the column amount of gas-phase HNO_3 . Some of this increase can be attributed to the evaporation of remaining nitric acid from the condensed phase as temperatures rise. This process is augmented by its three-body production from NO_2 and OH, and is in competition with destruction by photolysis



It is thought, however²¹, that the lifetime of HNO_3 against photolysis in the cold Antarctic stratosphere is many weeks. Changes in HNO_3 during late spring are also likely to be due to meridional mixing as the vortex begins to break up.

Long periods of cloud cover hindered data acquisition during the 1991–92 summer period. The limited data show, however, that nitric acid amounts at the end of December 1991 and in mid-January 1992 were $\sim 2.0 \times 10^{16}$ molecule cm^{-2} , consistent with the measurements for January 1991. The 1991 autumn measurements were made before the eruption of Mount Pinatubo, so we can discount the effects of heterogeneous chemistry due to an increased volcanic aerosol loading as observed at lower latitudes²⁰. By the end of September, however, there was evidence of volcanic aerosol in the polar vortex¹³.

The significance of the data lies in the way that odd nitrogen is partitioned in autumn in the long-lived reservoir of HNO_3 , rather than the temporary reservoir of N_2O_5 . In the Antarctic lower stratosphere, for solar zenith angles typical of late autumn and early spring, the lifetime of HNO_3 against photolysis is

many weeks, whereas that of N_2O_5 is a few days²¹. Conversion of NO_2 to HNO_3 therefore effectively removes NO_x from the atmosphere and thus its potential to lock up chlorine in a nonreactive form. We believe that the consistently low values of HNO_3 measured inside the vortex in spring (Fig. 2) result from conversion of HNO_3 from the gas to the solid phase, probably accompanied by loss from the atmosphere altogether by sedimentation. The very low values of NO_2 observed at the end of autumn and the beginning of spring imply that the NO_x reservoir is almost empty, with the remaining fraction lying at altitudes above the ozone depletion region. An accelerated loss of NO_2 through aerosol chemistry is therefore likely to increase the rate of ozone reduction in autumn, but have little effect in spring when ozone losses will be dominated by heterogeneous reactions on polar stratospheric clouds in the lower stratosphere. We therefore argue that aerosol chemistry in Antarctica will, through its added influence in autumn, contribute to both the extent and the duration of the annual ozone depletion.

The consequences at mid-latitudes of NO_x reduction through aerosol chemistry are difficult to assess. Any ozone depletion is likely to be small, however, because despite the huge increases in reactive surfaces provided by the Pinatubo aerosol, which resulted in a mid-latitude loss of 20% NO_2 , there was no apparent reduction in the ozone column²⁰.

In this analysis we have not attempted to quantify the role of chlorine chemistry in the partitioning of NO_x . Heterogeneous conversion of ClONO_2 and HCl to HNO_3 is likely to be unimportant on background aerosols¹¹, whereas conversion of N_2O_5 to HNO_3 by reaction (4) is likely to be important.

We conclude that the build-up of the HNO_3 column in autumn is consistent with N_2O_5 uptake on background aerosols, and that the loss of HNO_3 over winter is consistent with reactions

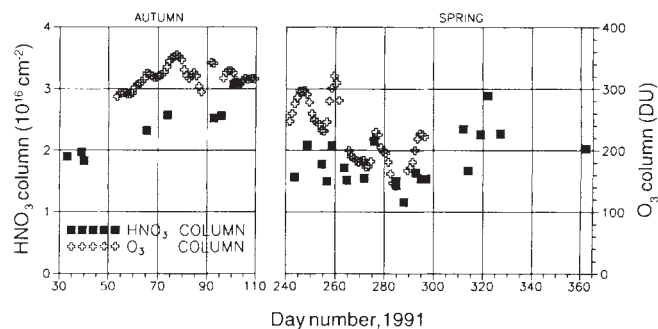


FIG. 2 Nitric acid vertical column amounts in units of molecules cm^{-2} derived from direct-Sun FTIR observations in the $866\text{--}870\text{ cm}^{-1}$ spectral region, and the approximate ozone column obtained from scattered sunlight measurements at twilight. We attribute the steady increase in nitric acid during autumn and the reduction during winter to the effects of heterogeneous chemistry. Wave-driven variations in the ozone column in spring can be seen to correlate with those in the HNO_3 data. We have taken the HNO_3 vertical profile for 65°S in May from the LIMS satellite results²², autumn temperatures from the May 80°S zonal mean climatology of Barnett and Corney²³, and springtime temperatures from balloonsonde flights¹³. Errors could be introduced by inaccurate assumptions in obtaining the vertical profiles of temperature, pressure and nitric acid and by uncertainties in the line parameters, although sensitivity tests of retrievals using LIMS profiles for January and May with appropriate temperature and pressure data show a change of only 1%. As a further test, we used a greatly perturbed nitric acid vertical profile¹⁵ for spring in which HNO_3 was removed between 15 and 20 km. Temperature data were taken from ref. 24. With this profile, HNO_3 column amounts are $\sim 17\%$ lower than with our 'standard' profile, so that an autumn/spring offset of this magnitude could conceivably result from inadequate knowledge of the actual HNO_3 vertical distribution. In general we estimate that sources of error in the retrievals limit the precision to about $\pm 10\%$ but that the absolute errors, taking into account uncertainties in the line parameters, are more likely to be of the order of $\pm 15\%$.

occurring on polar stratospheric clouds. Because of the importance of coupling between depleted NO_x , elevated ClO_x and ozone loss, we suggest that the data have significance in the wider global context, even though mid-latitude ozone loss through background aerosol chemistry has not been demonstrated directly. □

Received 1 May; accepted 25 November 1992.

1. Crutzen, P. & Arnold, F. *Nature* **324**, 651–655 (1986).
2. Solomon, S., Garcia, R. R., Rowland, F. S. & Wuebbles, D. *Nature* **321**, 755–758 (1986).
3. Rodriguez, J. M., Ko, M. K. & Sze, N. D. *Nature* **352**, 134–137 (1991).
4. Deshler, T. et al. *Geophys. Res. Lett.* **19**, 1819–1822 (1992).
5. Hofmann, D. J. et al. *Nature* **359**, 283–287 (1992).
6. Molina, L. T. & Molina, M. J. *J. phys. Chem.* **91**, 433–436 (1987).
7. Solomon, S. *Nature* **347**, 347–354 (1992).
8. Hofmann, D. & Solomon, S. *J. geophys. Res.* **94**, 5029–5041 (1989).
9. Solomon, S. & Keys, J. G. *J. geophys. Res.* **97**, 7971–7978 (1992).
10. Austin, J. A., Garcia, R. R., Russell, J. M. III, Solomon, S. & Tuck, A. F. *J. geophys. Res.* **91**, 5477–5485 (1986).
11. Hanson, D. R. & Ravishankara, A. R. *J. geophys. Res.* **96**, 5081–5090 (1991).
12. Reihls, C. M., Goldman, D. M. & Tolbert, M. A. *J. geophys. Res.* **95**, 16545–16550 (1990).
13. Johnson, B. J., Deshler, T. & Thompson, R. A. *Geophys. Res. Lett.* **19**, 1105–1108 (1992).
14. Fahey, D. W. et al. *Nature* **344**, 321–324 (1990).
15. Toon, G. C. et al. *J. geophys. Res.* **94**, 16571–16596 (1989).
16. Keys, J. G. & Johnston, P. V. *Geophys. Res. Lett.* **13**, 1260–1263 (1986).
17. Farmer, C. B., Toon, G. C., Schaper, P. W., Blavier, J.-F. & Lowes, L. L. *Nature* **329**, 126–130 (1987).
18. Mucray, F. J., Goldman, A., Blatherwick, R., Matthews, A. & Jones, N. *J. geophys. Res.* **94**, 16615–16618 (1989).
19. Coffey, M. T., Markin, W. G. & Goldman, A. *J. geophys. Res.* **94**, 16597–16613 (1989).
20. Johnston, P. V., McKenzie, R. L., Keys, J. G. & Matthews, W. A. *Geophys. Res. Lett.* **19**, 211–213 (1992).
21. Kawa, S. R. et al. *J. geophys. Res.* **97**, 7905–7923 (1992).
22. Russell, J. M. III et al. *Handbook for MAP* **22**, (1986).
23. Barnett, J. J. & Corney, M. *Handbook for MAP* **16**, 47–136 (1985).
24. Hofmann, D. J., Rosen, J. M., Harder, J. W. & Hereford, J. V. *J. geophys. Res.* **94**, 11253–11269 (1989).

ACKNOWLEDGEMENTS. We thank G. Avery, A. Thomas, M. Kotkamp and P. Kraak for operating the interferometer at Arrival Heights, S. David for help with data analysis and the New Zealand Antarctic Programme for logistic support.

Contribution of Late Permian palaeogeography in maintaining a temperate climate in Gondwana

Keddy Yemane*

Geological Institute, Swiss Federal Institute of Technology (ETH), Zürich, CH 8092, Switzerland

NUMERICAL simulations based on general circulation and energy-balance models consistently indicate that the high latitudes of Gondwana experienced seasonal extremes in climate during the Late Permian period^{1–3}. But palaeogeographic maps based on the distribution of climate-sensitive rocks⁴, palynological and palaeobotanical data^{5,6} and dicynodont fossil records⁷ all imply a temperate climate. The reason for this discrepancy has not been clear. Recently, it has emerged from studies of Upper Permian fluviolacustrine deposits throughout southern Africa that the geography was dominated by a series of giant lakes, perhaps interconnected within major fluvial frameworks⁸. Here I review these data and their implications for a temperate climate. I suggest that the discrepancy between the climate modelling results which indicate seasonal temperature extremes, and the increasing body of geological information documenting a temperate climate, may be explained by the fact that the palaeogeography used in the models does not take into account the existence of these lakes and rivers, which would have had a major influence on the regional climate. The results demonstrate the importance of incorporating accurate palaeogeographies into numerical modelling studies when attempting to reconstruct past climates.

* Present address: Department of Geology, University of Pennsylvania, 240 South 33rd Street, Philadelphia, Pennsylvania 19104-6316, USA.

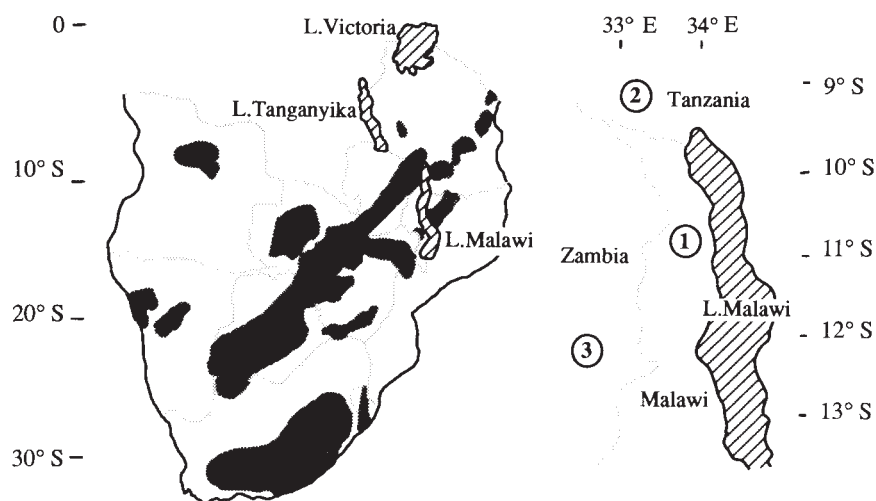
At the palaeolatitudes of southern Africa during the Late Permian, general circulation model (GCM)-based simulations show summer temperatures above 40 °C, winter temperatures below –30 °C, and surface temperature ranges of >50 °C (ref. 1). Community circulation (CCM)-based simulations indicate 60 °C summers and –42 °C winters². Energy balance model (EMB)-based simulations produce >45 °C summers and <–20 °C winters³. But palaeoclimate reconstructions for southern Gondwana based on evidence from sedimentary deposits indicate otherwise. The distribution of climate-sensitive rocks indicates that there was a humid belt at the palaeolatitudes of southern Africa⁴, and the palynofloral distribution indicates that it lay between palaeolatitudes 45° and 60°S (ref. 5). A map of plant distribution based on palaeobotanical data suggests that the Late Permian 'Glossopteris biome' was comparable to the modern 'cool temperate' Biome⁶, which consists of deciduous angiosperm forests in temperate high latitudes where average daily temperatures are >10 °C for at least one-third of the year⁹. Therapsid fossil distribution also suggests a temperate, moist climate at higher palaeolatitudes in Pangea⁷. The macro- and microflora from Upper Permian deposits in southern Africa are dominated by *Glossopteris*, which formed monospecific forests or woodlands along river banks and on elevated grounds¹⁰. The genus *Glossopteris* consisted of arborescent and herbaceous plants with roots adapted for growth in semiaquatic environments¹¹. Conspicuous growth rings in roots, trunks and branches, as well as periodic leaf falls, record the seasonal, frost-free growth patterns of deciduous plants^{12,13}.

The climate extremes suggested by the physical models would have rendered southern Africa inhospitable for both animal and plant life, yet fossils of the most diverse and prolific terrestrial vertebrates, the mammal-like reptiles dicynodonts, are found in South Africa, Zambia, Zimbabwe, Malawi and Tanzania¹⁴. These mouse- to bear-sized herbivores subsisted on *Glossopteris* foliage, ferns, low-lying conifers and roots in flood plain and lake shore environments¹⁵. It has been suggested³ that seasonal migrations could explain their presence, but this hypothesis is untenable. Dicynodonts, with their small body size, sprawling reptilian gait and lack of internal thermoregulatory mechanisms¹⁶, would have been unable to migrate over long distances. The geological record does not suggest extended migration of dicynodonts or other Late Permian therapsids; sites of dicynodont remains within Gondwana are primarily restricted to southern Africa^{7,16}. Distant sites in Asia and Europe where comparable faunal remains have been recorded¹⁶ are unlikely seasonal refuges for dicynodont herds. The presence of contemporary amphibians, such as the labyrinthodonts¹⁷, shows that temperatures must have been moderate and water bodies present year-round, as migration would not have been an option. Although unusual semi-fossorial habits have been reported (for example *Diictodon*)^{18,19}, there is no definitive evidence for underground aestivation or hibernation among dicynodonts. Rather than invoking preservation bias or seasonal occupation of the habitat to explain the geological record, we should rule out the freezing winters and torrid summers simulated by the climate models^{1–3}.

Palaeoclimatic signatures preserved in Upper Permian sedimentary sequences from southern Africa corroborate those from the fossil record. The known extent of the fluvial and lacustrine deposits (Fig. 1) suggests the existence of concurrent, possibly interconnected, giant lakes throughout the subcontinent⁸. Whereas fluvial-dominated deposits formed in foreland basin settings²⁰ at the southern edge, lacustrine-dominated deposits accumulated in settings of low relief, broad warping and mild faulting²¹ at the northern end. These deposits, thousands of miles inland from the nearest sea, indicate a period during which vast lakes developed. I will next evaluate the climatic conditions needed to sustain this uniquely wet environment at the heart of a supercontinent.

The dominantly clastic (shales) and lesser chemical

FIG. 1 Sketch of known surface and subsurface extent of Upper Permian fluviolacustrine deposits in southern Africa. See ref. 8 for source. (For comparison, Lake Victoria, the largest lake in Africa, has a water surface of 68,000 km², roughly the area of Ireland.) Inset: location of Upper Permian sequences referred to in Table 1. (1) Karonga and Mt Chombe, Chiweta (northern Malawi), (2) Tukuyu, Illima and Ngana Hills (southern Tanzania), (3) Luwanga Valley, northern Zambia.



(carbonates) deposits in northern Malawi, southern Tanzania and northern Zambia⁸ suggest that deposition took place in shallow lake basins which were relatively hydrologically open and in deeper lake basins which were relatively hydrologically closed. Minor siltstones, subarkoses and arkoses are interbedded within grey, greenish and black shales. The unbioturbated shales contain rhythmites including seasonal varves. The clay mineral assemblage predominantly constitute illite and chlorite; smectite is absent or occurs in minor quantities as interlayers. The pollen flora is dominated by striated bisaccates and polyapiculates, and monosaccates which correlate with the *Striatiti* pollen zone²² of Permo-Triassic (Lower Beaufort) sequences in southern Africa. Similar sedimentological attributes are recorded from equivalent sequences throughout southern Africa⁸ and even from further south in Antarctica (~80–85° S palaeolatitudes)²³. Sedimentary features in modern lacustrine depositional settings comparable to those found in the Upper Permian sequences are summarized in Table 1. The integrated climate record of the biotic and sedimentological data suggests cool, temperate, humid, seasonal climate during Upper Permian deposition in southern Africa (~55–60° S) and beyond.

The areal extent of Upper Permian deposits in southern Africa alone, not to mention in Gondwana in general, cover several

million square kilometres. Such depositional settings could not have existed under the temperature extremes and humidity (annual precipitation minus evaporation of 0–1 mm day⁻¹)¹ simulated by the models. The unique Late Permian palaeogeography does not feature in any of the current physical climate models, usually on the grounds that dimensions of unit cells used in sensitivity experiments are too large to include such details, or that the geography is a result of 'microclimate' influences. Both of these arguments are invalidated by the geographical extent of fluviolacustrine deposits in southern Gondwana.

To justify the simulated climate extremes, the extreme seasonal temperatures currently experienced in Canada and Siberia have been invoked³. The high latitudes, low-lying continental interiors, cratonic tectonic setting, predominant metamorphic terrains, coniferous vegetation stands and extensive lakes may indeed be comparable to the Late Permian palaeoenvironments, but the climate-responsive sedimentary record from the lake deposits demonstrates that the analogy ends there. Summer and winter temperatures in the Canadian interior are generally 0–22 °C and –10–28 °C, respectively, but summer maxima of >40 °C and winter minima of <–40 °C have been recorded locally²⁴. The northern Great Plains of western Canada (50°–60° N) contain more than one million fresh, brackish and saline

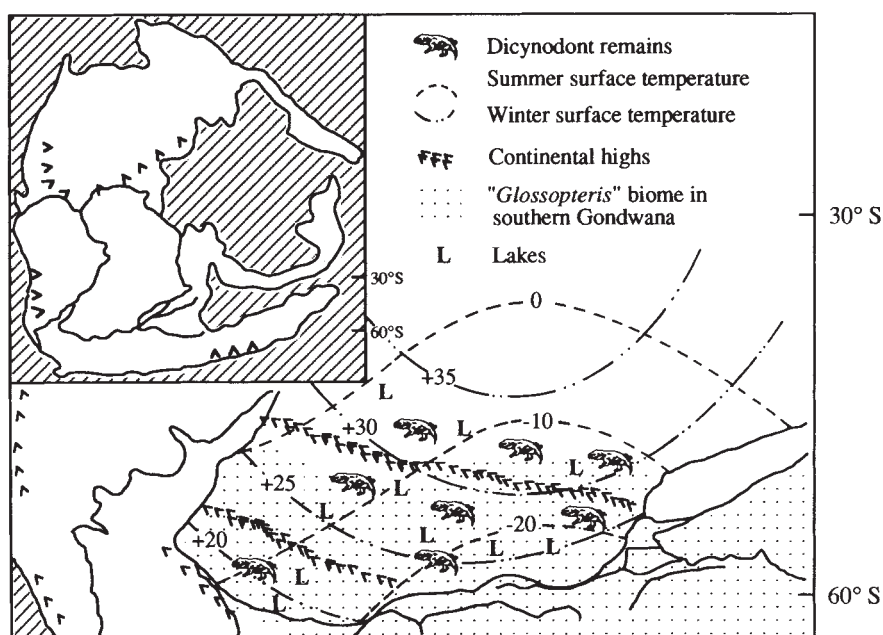


FIG. 2 Palaeoenvironmental reconstruction for southern Africa during the Late Permian. Shown are palaeogeography⁴⁸, geomorphological features during late Early Permian time⁴⁹, vegetation⁶ and some major known sites of dicynodont remains. Note isotherms for numerically simulated summer and winter surface temperatures^{1,3} (in °C).

TABLE 1 Indications of temperate climate

Stratigraphy and mineralogy	The Upper Permian-age Mt Chombe sequence from northern Malawi is underlain by Lower Permian coal measures with associated mudstones and sandstones and is overlain by the lacustrine chiweta bone beds, comprising mudstones and limestones with bone fragments and dicynodont skulls ³⁴ . Angular to subangular quartz derived mostly from low-grade metamorphic rocks dominates the sediments ³⁵ . Fresh, angular sodic plagioclase and orthoclase constitute the second most abundant mineral group, but small quantities of microcline and biotite to occur. Similar facies and clastic mineralogy are observed in the stratigraphically equivalent Ruhuhu beds, Tanzania ³⁶ and Madumabisa mudstones, Zambia (work in progress).
Key sedimentary features	Annual varves from Mt Chombe sequence comprise alternating low-Mg calcite, amorphous organic matter and clay laminae. Fine laminations comprising low-Mg calcite silt, silt/organic matter, clay/silt, clay/organic matter and silt/sand laminae are abundant. Couplet comprising clay/calcite/silt laminae and representing 6,000 years of deposition occur ³² in an equivalent sequence from northeast Angola. Nonglacial varves with couplets comprising calcite/diatomite/organic matter/clay laminae currently form in temperate lakes for example Lake Zürich ³⁸ . Pedogenic features such as palaeosols, desiccation cracks, red coloration and <i>in situ</i> plant growth are absent from the northern Malawi sequence ³⁵ .
Clay mineral assemblages	Two detrital clay assemblages, illite-chlorite-mixed layers (more seasonal) and chlorite-illite (less seasonal) assemblages, corresponding to three cycles are documented from the northern Malawi sequence. Regular and irregular mixed layers with open illites dominate during periods with seasonal aridity whereas chlorite-dominated assemblages with tightly closed illite prevail during wetter periods. Authigenic clays are insignificant. Illite, mixed layers, chlorite and smectite are reported from the Ruhuhu Beds, Tanzania ³⁶ . Preliminary data from Madumabisa Mudstones in the Luwanga Valley, Zambia, show illite and mixed layers as the main constituents. Modern high-latitude temperate lakes have quantitatively comparable major detrital clay mineral assemblages comprising illite, chlorite and mixed layers ³⁹⁻⁴¹ .
Micro- and Macroflora assemblages	Microflora from northern Malawi sequence comprises (>90%) <i>Striatopodocarpites</i> , <i>Vittatina</i> , <i>Densipollenites</i> and <i>Gutullapollenites</i> . <i>Taeniasporites</i> and <i>Leuckisporites</i> are rare. Spores are under-represented (<5%). Plant macrofossils such as <i>Glossopteris browniana</i> , <i>G. Indica</i> , <i>Vertebraria</i> and <i>Schizoneura</i> were also reported from northwestern Malawi ⁴² . Pollen assemblages from Zambia ⁴³ , Zimbabwe ⁴⁴ and Tanzania ³⁶ have comparable floral composition.
Oxygen isotope signatures	¹⁸ O signatures were measured on early-formed low-Mg calcite cements and septarian vein fills in near-spherical concretions from the Chombe sequence and from equivalent sequences in southern Tanzania. $\delta^{18}\text{O}$ ranges between -11 and -13‰ PDB in concretion cements, and between 12 and -14‰ PDB in vein calcites. Modern meteoric precipitations ⁴⁵ and regional ground waters ⁴⁶ at equivalent high latitudes have lighter isotope signatures $\delta^{18}\text{O}$. Palaeotemperatures calculated from $\delta^{18}\text{O}$ of the meteoric concretion cements and vein-fills ⁴⁷ suggest mean annual air temperatures of ~10 °C. Similar air temperatures are observed today in geographically equivalent temperate high latitudes.

Summary of attributes suggesting seasonal cool temperature climate during Late Permian deposition in southern Africa⁴⁸. Data were gathered from Upper Permian lacustrine sequences in northern Malawi, southern Tanzania and northern Zambia. See Fig. 1 for location.

lakes. Laterally continuous thick beds of salts (mirabilite, thenardite and epsomite) are mined from the saline lake deposits^{25,26}. In the intramontane lakes of the Cariboo plateau in British Columbia, thick salts (natron, epsomite, bloedite and mirabilite) form interstitially within lake beds²⁷. The predominant clay mineral constituents in lake muds from both regions include smectite, illite-smectites and lesser kaolinite. The west Siberian plain (50°–60° N) where average summer and winter temperatures range from 3 to 23 °C and -16° to -28 °C, respectively²⁸, contain more than one hundred fresh, brackish and saline lakes. Thick organic-rich evaporites of salts (halite, mirabilite and thenardite) are mined from shallow lakes in West and southeastern Siberia^{29,30}. The region is characterized by permafrost³¹, and smectite and kaolinite are the common authigenic clay minerals in the soils³².

In contrast, evaporite deposits or permafrost structures are not known from Upper Permian sequences in southern Africa. The biotic record and the sedimentary history of Upper Permian sequences throughout southern Africa and beyond document palaeogeographic settings unparalleled in modern times. The extensive fluvial environments with freshwater lakes, lush vegetation and thriving fauna (Fig. 2) could have existed only in a moderate climate conditions with adequate humidity. One means suggested for transfer of moisture into the interior of Pangaea has been monsoonal circulation created by the Permian continental configuration^{4,7}. Seasonal rains drenching uplands may have contributed to the water budget, but the annual precipitation of <2 mm day⁻¹ simulated for high-latitude Gondwana¹ could not have sustained the vast perennial freshwater lakes in southern Africa. Winds blowing over lakes pick up a tremendous amount of moisture which is then precipitated downwind of the lakes. This 'lake effect' modifies the climate considerably within 40–50 km of individual lakes as is observed in the Great Lakes region of North America³³. By inference, the giant perennial lakes in southern Gondwana may have moderated the regional climate during the Late Permian. The fact that the numerical models do not incorporate this unique wet Late Permian palaeogeography could well account for the uniformly erroneous results.

The sedimentary records of lake and associated fluvial deposits should be used to define realistic boundary conditions for palaeoclimate simulations. This is particularly true for most of Carboniferous–Jurassic Gondwana where the geological history is essentially documented in the records of continental deposits. Thus, the appropriate geological information, together with cell dimensions that would accommodate even smaller continental deposits, is necessary for sensitivity experiments for the Late Permian or other critical time slices. If the aim of palaeoclimate modelling is to 'recreate' past climates, the geological evidence must be given more weight than it is now. Agreement between models, which has been construed as showing the accuracy of simulated Late Permian climates, may only reflect the fundamental shortcomings in parameters employed by all models, such as the palaeogeography. Earth scientists should not be quick to modify geological data to accommodate model simulations; rather, the acceptability of climate simulations should be judged on how well they reproduce independently established geological facts. This would strengthen the theoretical foundations for numerical modelling, thus permitting accurate reconstructions of the past and realistic predictions of future global climate trends. □

Received 12 August; accepted 17 November 1992.

1. Kutzbach, J. E. & Gallimore, R. G. *J. geophys. Res.* **94**, 3341–3357 (1989).
2. Moore, G. T., Hayashida, D. N., Ross, C. A. & Jacobson, S. R. in *Abstr. Vol. Project Pangaea, May 23–27, 1992* (Kansas, 1992).
3. Crowley, T. J., Hyde, W. T. & Short, D. A. *Geology* **17**, 457–460 (1989).
4. Robinson, P. L. in *Implications of Continental Drift to the Earth Sciences* (ed Tarling, D. H. & Runcorn, S. K.) 451–476 (Academic, London, 1973).
5. Kremp, G. O. W. *Paleo-Data-Banks* **7**, 1–21 (1977).

6. Ziegler, A. M. in *Palaeozoic Palaeogeography and Biogeography* (eds McKerron, W. S. & Scotese, C. R.) 363–379 (Geological Society Memoir, Oxford, 1990).
7. Parrish, J. M., Parrish, J. T. & Ziegler, A. M. in *The Ecology and Biology of Mammal-like Reptiles* (eds Maclean, P. D., Roth, J. J. & Roth, E. C.) 109–131 (Smithsonian Institution Press, Washington DC, 1986).
8. Yemane, K. & Kelts, K. in *Major African Continental Phanerozoic Complexes and Dynamics of Sedimentation* (eds Kogbe, C. & Lang, J.) 169–185 (Pergamon, Paris, 1990).
9. Walter, H. *Vegetation of the Earth*, 3rd Edn (Springer, Berlin, 1985).
10. Anderson, J. M. *Mem. Bot. Surv. S. Africa* **41** (1977).
11. Stewart, W. N. *Paleobotany and the Evolution of Plants* (Cambridge Univ. Press, Cambridge, 1983).
12. Plumstead, E. P. in *Atlas of Palaeobiogeography* (ed. Hallam, A.) 187–205 (Elsevier, Amsterdam, 1973).
13. Taylor, E. L., Taylor, T. N. & Cúneo, N. R. *Science* **257**, 1675–1677 (1992).
14. Kitching, J. W. *Mem. Bernard Price Inst. Palaeont. Res.* **1**, 1–131 (1977).
15. Hotton, N. III, in *The Ecology and Biology of Mammal-like Reptiles* (eds Maclean, P. D., Roth, J. J. & Roth, E. C.) 71–82 (Smithsonian Institution Press, Washington DC, 1986).
16. King, G. in *The Dicotyledons, A Study in Palaeobiology* (Chapman and Hall, London, 1990).
17. Colbert, E. H. in *The Ecology and Biology of Mammal-like Reptiles* (eds Maclean, P. D., Roth, J. J. & Roth, E. C.) 133–145 (Smithsonian Institution Press, Washington DC, 1986).
18. Cluver, M. A. *Ann. S. Afr. Mus.* **76**, 213–246 (1978).
19. Smith, R. M. H. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **60**, 155–170 (1987).
20. Tankard, A. J. *et al.* *Crustal Evolution of Southern Africa: 3.8 byr of Earth History* (Springer, New York, 1982).
21. Rust, I. in *Gondwana Geology* (ed. Campbell, K. S. W.) 537–564 (Australian National Univ. Press, Canberra, 1975).
22. Hart, G. F. in *Aspects of Palynology* (eds Tschudy, R. H. & Scott, R. A.) 271–289 (Wiley, New York, 1969).
23. Isbell, J. L. thesis, Ohio State Univ. (1990).
24. *Climate Atlas of Canada: A Series of Maps Portraying Canada's Climate. Canadian Climate Program Map Series 1—Temperature and degree days* (Canadian Government Publishing Center, Ottawa, 1984).
25. Last, W. M. *Sedim. Geol.* **64**, 207–221 (1989).
26. Teller, J. T. & Last, W. M. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **76**, 215–240 (1990).
27. Renaut, R. W. & Long, P. R. *Sedim. Geol.* **64**, 239–264 (1989).
28. *Encyclopaedia Britannica, Macropaedia* **28**, (Chicago, 1991).
29. Trofimov, V. T., Kudryashov, V. G. & Galaktionov, B. V. *Sov. Engng Geol.* **1**, 63–86 (1989).
30. Stankevich, E. F., Batalin, Y. V. & Sinyavskii, E. I. *Sov. Engng Geol.* **31**, 31–37 (1990).
31. Melnikov, V. P., Devyatkin, V. N. & Bevzenko, Y. P. in *Proc. 5th int. Congr. Permafrost*, Vol. 1 (ed. Senneset, K.) 815–818 (Trondheim, 1988).
32. Gradusov, B. P. & Sokolov, I. A. *Sov. Soil Sci.* **22**, 80–87 (1990).
33. Eichenlaub, V. *The Weather and Climate of the Great Lakes Regions* (Univ. of Notre Dame Press, 1979).
34. Dixey, F. *Trans. geol. Soc. S. Africa* **29**, 59–68 (1972).
35. Yemane, K., Siegenthaler, C. & Kelts, K. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **70**, 165–178 (1989).
36. Kreuser, T. *Geol. Inst. Univ. Köln Sonderveröffentlichungen* **45**, 1–217 (1983).
37. Oesterlen, M. *Geol. Jb. Riehe* **B36**, 1–44 (1979).
38. Kelts, K. & Hsu, K. in *Lakes, Chemistry, Geology, Physics* (ed. Lerman, A.) 295–323 (Springer, New York, 1978).
39. Englund, J. O., Jørgensen, P., Rødsdøl, E. & Aagaard, P. in *Proc. int. Symp. Interactions between Sediments and Freshwater* (ed. Golterman, H. L.) 125–132 (The Hague, Netherlands, 1977).
40. Muller, G. & Quakernaat, J. *Contr. Mineral. Petrol.* **22**, 268–275 (1969).
41. Moore, J. E. *J. sedim. Petrol.* **31**, 402–436 (1961).
42. Andrew, A. R. & Bailey, T. E. G. *Q. J. geol. Soc. Lond.* **66**, 202–204 (1910).
43. Utting, J. in *Proc. 4th int. Palynology Conf. Lucknow, India, 1976* (eds Bharadway, D. C., Singh, H. P. & Tiwari, T. S.) 208–219 (Tej Kumar, Lucknow, 1980).
44. Falcon, R. in *Proc. 4th int. Palynology Conf. Lucknow, India, 1976* (eds Bharadway, D. C., Singh, H. P. & Tiwari, T. S.) 165–174 (Tej Kumar, Lucknow, 1980).
45. Yurtsever, Y. & Gat, J. R. in *The Water Cycle* (eds Gat, J. R. & Gonfiantini, R.) 103–142 (International Atomic Energy Agency, Vienna, 1981).
46. Gat, J. R. *Water Resources Res.* **7**, 980–993.
47. Hays, P. D. & Grossman, E. L. *Geology* **19**, 441–444 (1991).
48. Yemane, K. thesis, ETH (1990).
49. Scotese, C. R. & Golonka, J. *Paleogeographic Atlas, Paleomap Progress Report* no. 20 (Univ. of Texas at Arlington, 1992).
50. Visser, J. N. J. in *C. R. XII int. Congr. Carboniferous and Permian Stratigraphy and Geology* (Buenos Aires, Argentina, in the press).

ACKNOWLEDGEMENTS. I thank K. Kelts and K. Hsu for research guidance and ideas, P. Abbott, G. Boyajan, G. Fonjweng, R. Giegengack, I. Harker, A. Kihm, H. Pfefferkorn, N. Shubin, J. N. J. Visser, J. Westgate and M. Whalen for commenting on earlier versions of the manuscript, G. de V. Klein for a travel grant to attend the 1992 Pangaea Workshop, and A. Truneh for encouragement and support. This study was supported by Swiss Lake Research. The manuscript was prepared at the Geology Department, University of Pennsylvania.

Three-dimensional fluctuation conductivity in superconducting single crystal K_3C_{60} and Rb_3C_{60}

X.-D. Xiang, J. G. Hou, Vincent H. Crespi, A. Zettl & Marvin L. Cohen

Department of Physics, University of California at Berkeley, and Materials Sciences Division, Lawrence Berkeley Laboratory, Berkeley, California 94720, USA

THE superconducting transition temperature, T_c , defines the point at which the free energies of the superconducting and normal states of a material become equal. Just above T_c , thermodynamic fluctuations produce small, transient regions of the superconducting state, giving rise to an anomalous increase in the normal-state conductivity known as paraconductivity. This situation is analogous to the fluctuating regions of correlated spins found near the Curie–Weiss transition in ferromagnets. Such fluctuations are of theoretical significance in that they provide a direct probe of critical phenomena in general, and a stringent test of scaling theories, which describe the approach to the critical point. Paraconductivity effects are strongly dependent on the dimensionality of the system, although for conventional superconductors, three-dimensional fluctuation conductivity has to our knowledge never been observed. Here we report the observation of pure, three-dimensional paraconductivity in single crystals of the recently discovered¹ superconductors K_3C_{60} and Rb_3C_{60} . In addition to probing the critical state near T_c , these measurements allow the indirect determination of the residual, normal-state resistivity.

High-quality single-crystal specimens of K_3C_{60} and Rb_3C_{60} were prepared by doping vapour-transport-grown C_{60} crystals with alkali metals. The synthesis followed a method similar to that described previously² for K_3C_{60} . We used the standard in-line four-probe contact configuration and kept the d.c. probing current small (10–100 μA) to minimize Joule heating effects.

The insets to Fig. 1a and b show, for K_3C_{60} and Rb_3C_{60} , respectively, the (normalized) resistivity $\rho(T)$ as a function of

temperature over an extended temperature range. As demonstrated previously² for K_3C_{60} , the temperature dependence of ρ for Rb_3C_{60} is metallic. This behaviour is in contrast to the nonmetallic temperature dependence of $\rho(T)$ found for Rb-doped C_{60} when C_{60} crystals grown from a CS_2 solution were used³. Although the origin of this discrepancy is unknown, a nonmetallic $\rho(T)$ may result from incomplete doping or an impure host crystal. Our resistively determined T_c values (transition midpoint) for K_3C_{60} and Rb_3C_{60} are 19.8 K and 30.2 K, respectively. We note that the overall functional form for $\rho(T)$, except at temperatures below 50 K, is similar for K_3C_{60} and Rb_3C_{60} . The temperature dependence of $\rho(T)$ is reproducible over many samples in both cases.

Figure 1a and b also shows $\rho(T)$ for K_3C_{60} and Rb_3C_{60} in greater detail near T_c (the temperature scale has been normalized to T_c in each case). Just above T_c , the resistivity for both materials deviates from its normal-state behaviour before dropping precipitously at T_c . We associate this deviation with superconducting fluctuations. It should be recognized that identification of fluctuation effects of small magnitude requires data with a high signal-to-noise ratio. In an earlier investigation² of the resistivity of K_3C_{60} , fluctuation effects near T_c were not resolved because of large scatter in the data.

The paraconductivity, σ' , is obtained by subtracting the extrapolated normal-state conductivity σ_n from the measured conductivity. To establish an appropriate baseline for the paraconductivity, we have tried various power-law extrapolations using the temperature range $1.4T_c < T < 2T_c$ to establish fitting parameters for the extrapolation. For both materials, a linear $-T$ curve provides an adequate fit (solid line in Fig. 1a and b). In K_3C_{60} , a T^5 curve (dashed line in Fig. 1a) provides a slightly better fit at higher temperatures ($T \approx 60$ K). But as shown in Fig. 1a, the difference between the two extrapolations is small in the relevant temperature range near T_c , and both are consistent with ρ measured near T_c at high magnetic field (also shown in Fig. 1a). Without further physical basis, in our analysis we shall use the T^5 extrapolation for K_3C_{60} and the linear $-T$ extrapolation for Rb_3C_{60} . Our conclusions of three-dimensional paraconductivity are not sensitive to which extrapolation is used. The baseline choice does affect the determined numerical value of the residual normal-state resistivity (see below), but different

reasonable extrapolations lead to resistivities all within 20% of one another.

As first derived by Aslamazov and Larkin (AL)⁴, the excess conductivity σ' is given by⁵

$$\sigma'_{AL} \sim t^{-(4-D)/2} \quad (1)$$

where D is the dimensionality of the specimen and t is the reduced temperature defined by $t = (T - T_c)/T_c$. This AL term is known as the regular fluctuation conductivity, which is due to the direct acceleration of the fluctuation-induced superconducting pairs of quasiparticles. In addition to the AL term, Maki⁶ and Thompson⁷ (MT) have derived another contribution known as the anomalous fluctuation conductivity. The MT term results from the scattering of the normal quasiparticles by the superconducting fluctuations and is small when 'pair-breaking' effects are large.⁸ In three dimensions, the excess conductivity may be expressed as⁸

$$\sigma'_{3D} = \sigma'_{AL} + \sigma'_{MT} = \sigma_{exc} \left[t^{-1/2} + \frac{4}{t^{1/2} + \delta^{1/2}} \right] \quad (2)$$

where δ is a pair-breaking parameter. The first term on the right is the AL term, and the second is the MT term. σ_{exc} is a prefactor to be defined below.

Figure 2a and b shows on log-log plots the experimentally determined σ' against T/T_c for K_3C_{60} and Rb_3C_{60} . The insets show log-log plots of σ' against t for the same data sets. The paraconductivity data for K_3C_{60} and Rb_3C_{60} can be well accounted for by assuming an effective dimensionality $D = 3$. For K_3C_{60} (Fig. 2a), we find that including a MT contribution improves the fit (although the AL term dominates). From our best fit, we obtain a pair-breaking parameter $\delta \approx 0.58$ for K_3C_{60} . For Rb_3C_{60} (Fig. 2b), we find that a good fit is obtained only if the MT term is negligible, which implies a large δ . This is consistent with Rb_3C_{60} having a stronger electron-phonon coupling constant (resulting in a higher T_c) than K_3C_{60} (ref. 9). Only for $D = 3$ do we obtain good fits to the paraconductivity data of K_3C_{60} and Rb_3C_{60} . The inset to Fig. 2b shows how the data of Rb_3C_{60} compare to paraconductivity fits with $D = 1$ and $D = 2$ in addition to $D = 3$.

We observe no evidence for granular superconductivity. The coherence volume of the superconducting order parameter grows as the temperature approaches T_c unless limited by reduced dimensionality. For a granular superconductor with grain size of L , the functional form of the excess conductivity crosses over from three-dimensional behaviour to zero-dimensional behaviour when the Ginzburg-Landau coherence length $\xi(T)$ exceeds $L/3$ as T approaches T_c (ref. 10). In the

presence of 0-D crossover, the data of Fig. 2a and b would deviate upwards away from the theoretical curves based on 3-D fluctuations with decreasing reduced temperature t , and there would be a region of steeper slope (-2) at small t . Such behaviour is not observed for either K_3C_{60} or Rb_3C_{60} (see insets to Fig. 2a and b). The absence of a 0-D crossover down to $t \approx 0.0005$ in our data gives a lower limit for the domain size of about $0.6 \mu m$. This dimension is at least 100 times larger than the measured values of $\xi(0)$ ^{11,12}, and therefore guarantees that physical parameters measured on these samples reflect intrinsic properties.

An important parameter in describing electronic transport in K_3C_{60} and Rb_3C_{60} is the absolute magnitude of the resistivity. Because of uncertainties in sample dimensions, contact position and current distribution profiles (if a van der Pauw configuration is applied), it is difficult to make an accurate direct determination of the resistivity. Experimentally, we observe identical normalized fluctuation resistivity curves (therefore, identical σ'/σ_n) in different samples with somewhat different estimates of $\rho(0)$ based on direct measurements. However, the normal-state conductivity σ_n may be determined indirectly through the measured normalized conductivity (σ'/σ_n), because the excess conductivity prefactor σ_{exc} is related to the independently determined coherence length $\xi(0)$ by

$$\sigma_{exc} = \frac{e^2}{32\hbar\xi(0)} \quad (3)$$

Using equation (2) together with the experimentally determined values (σ_{exc}/σ_n)_K $\approx 2 \times 10^3$, (σ_{exc}/σ_n)_{Rb} $\approx 7.3 \times 10^{-3}$, $\xi(0)_K \approx 45 \text{ \AA}$ (ref. 11) and $\xi(0)_{Rb} \approx 24 \text{ \AA}$ (ref. 12), we obtain the zero-temperature residual normal-state resistivities $\rho(0)_K \approx 0.12 \text{ m}\Omega \text{ cm}$ and $\rho(0)_{Rb} \approx 0.23 \text{ m}\Omega \text{ cm}$, where the subscripts K and Rb represent K_3C_{60} and Rb_3C_{60} , respectively. The different extrapolations of the normal-state resistivity produce an uncertainty of about $\pm 20\%$ in $\rho(0)$. The resistivity values are reasonably consistent with our most reliable direct measurement on K_3C_{60} ($\sim 0.5 \text{ m}\Omega \text{ cm}$) and with infrared studies¹³ on Rb_3C_{60} ($\sim 0.4 \text{ m}\Omega \text{ cm}$). The results also agree with our theoretical calculation¹¹ ($\rho(0)_K \approx 0.2 \text{ m}\Omega \text{ cm}$) based on H_{c2} measurements and a theoretical calculation⁴ ($\rho(0)_K \approx \rho(0)_{Rb} \approx 0.39 \text{ m}\Omega \text{ cm}$) in which a maximum disorder (defined as equal distribution of two possible orientations of C_{60} in a face-centred cubic structure) was assumed for both K_3C_{60} and Rb_3C_{60} . Our results suggest that Rb_3C_{60} may have a more C_{60} orientational disorder than K_3C_{60} .

As this is, to our knowledge, the first observation of pure 3-D paraconductivity in any isotropic superconductor, it seems appropriate to address the difficulty of observing similar

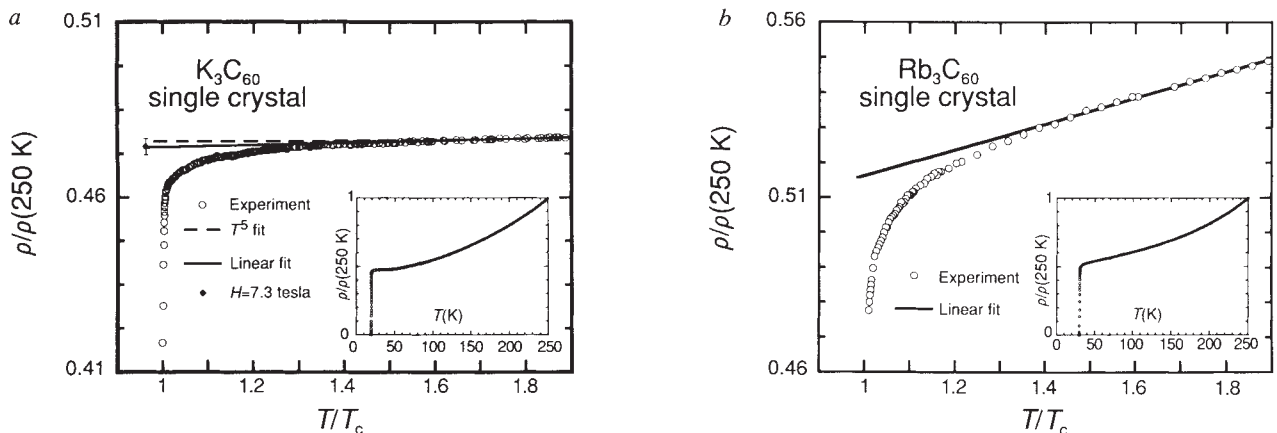


FIG. 1 Normalized resistivities as a function of T/T_c for K_3C_{60} (a) and Rb_3C_{60} (b). Circles are experimental data and solid lines are linear extrapolations to the normal-state resistivity. For K_3C_{60} , a T^5 extrapolation is also

shown along with a high-magnetic-field resistivity point. Insets: normalized resistivities against temperature for K_3C_{60} (a) and Rb_3C_{60} (b).

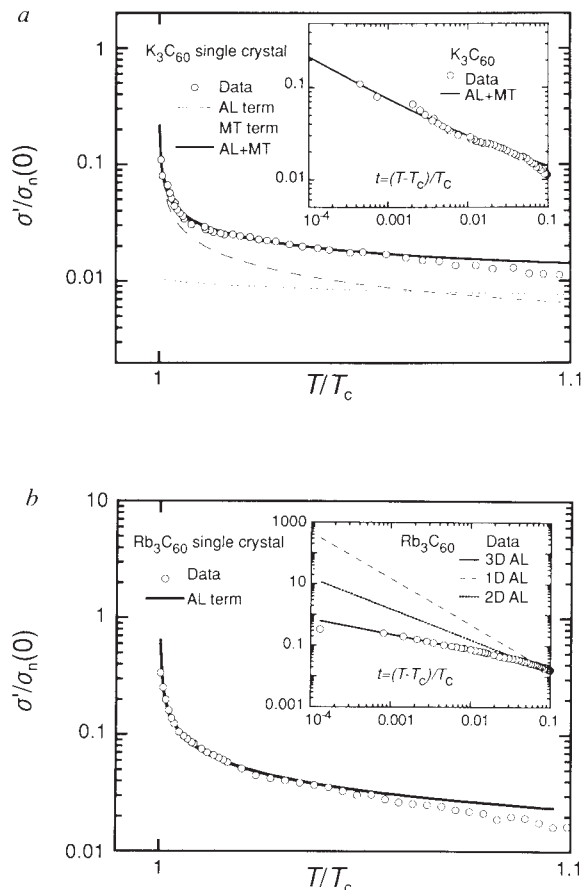


FIG. 2 Log-log plots of normalized fluctuation conductivities against T/T_c for K_3C_{60} (a) and Rb_3C_{60} (b). Insets: log-log plots of normalized fluctuation conductivities against reduced temperature $t = (T - T_c)/T_c$ for K_3C_{60} (a) and Rb_3C_{60} (b).

behaviour in other materials (most conventional superconductors are, of course, three-dimensional). For conventional 3-D superconductors, the fractional change in normal conductivity $\sigma'/\sigma_n \approx (kT_c/E_F)(1/k_F l)(T/(T - T_c))^{1/2}$ is of order of $10^{-7}(T/(T - T_c))^{1/2}$, too small to be observable in any meaningful temperature interval above T_c . Here E_F is the Fermi energy, k_F the Fermi wave vector, and l the mean free path. To observe 3-D fluctuation conductivity, large concentrations of impurities or defects have to be induced, and these are usually accompanied by granularity or reduced dimensionality. For anisotropic-layered high- T_c materials, the situation is more complicated. Paraconductivity has been reported for many different compounds in various forms, but no consensus about the dimensionality in these materials has been reached^{8,15}. In the fullerene-based superconductors, the intrinsic orientational disorder (leading to a relatively high intrinsic resistivity or short mean free path), short coherence length and high T_c together greatly extend the useful temperature range for measurement of the fluctuations and allow direct observation of the fluctuation phenomena in pure single crystals. Our results demonstrate that these systems are genuinely 3-D superconductors as opposed to lower-dimensional or granular superconductors. □

6. Maki, K. *Prog. theor. Phys.* **39**, 887 (1968); **40**, 193 (1968).
7. Thompson, R. S. *Phys. Rev. B* **1**, 327 (1970); *Physica* **55**, 296 (1971).
8. Maki, K. & Thompson, R. S. *Phys. Rev. B* **39**, 2767 (1989).
9. Appel, J. *Phys. Rev. Lett.* **21**, 1164 (1968).
10. Schmidt, H. Z. *Phys.* **216**, 336 (1968).
11. Hou, J. G., Crespi, V. H., Xiang, X.-D., Zettl, A. & Cohen, M. L. *Phys. Rev. Lett.* (submitted).
12. Sparr, G. et al. *Phys. Rev. Lett.* **68**, 1228 (1991).
13. Rotter, L. D. et al. *Nature* **355**, 532 (1992).
14. Gelfand, M. P. & Lu, J. P. *Phys. Rev. B* **46**, 4367 (1992).
15. Reggiani, L. & Vaglio, R. *Phys. Rev. B* **44**, 9541 (1991).

ACKNOWLEDGEMENTS. We thank W. A. Vareka, G. Briceno and B. Burk for technical assistance. This research was supported by the NSF and by the Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division of U.S. Department of Energy. V.H.C. was supported by a National Defense Science and Engineering Graduate Fellowship.

Palaeomagnetic evidence for large-magnitude, low-angle normal faulting in a metamorphic core complex

Richard F. Livaccari*, John W. Geissman* & Stephen J. Reynolds†

* Department of Earth and Planetary Sciences, University of New Mexico, Albuquerque, New Mexico 87131-1116, USA

† Department of Geology, Arizona State University, Tempe, Arizona 85287, USA

CONTROVERSY exists over whether large-magnitude extensional deformation can occur along low-angle, master detachment (normal) faults¹⁻¹⁶. Large amounts of Cenozoic extension occurred in the Basin and Range province of the western United States, and master detachment faults related to this extension are currently exposed as subhorizontal structures along crustal arches known as metamorphic core complexes (MCCs)^{10,11,17}. One set of models suggests that MCCs represent crustal-scale blocks bounded by originally high-angle normal faults which have tilted to subhorizontal attitudes¹⁻⁶. More dynamic models propose isostatically induced flexural tilting of an initial subhorizontal fault, to an active high-angle fault, to final abandonment as a subhorizontal structure⁷⁻⁹; these models require MCC detachment faults and their footwalls to have been tilted by 30-60°. Alternatively, MCC detachment faults may have originated as subhorizontal or low-angle structures¹⁰⁻¹⁶. Here we test these models with palaeomagnetic data from undeformed portions of the South Mountains, a typical Cenozoic MCC in the southern Basin and Range province¹⁸⁻²⁴. Comparison with time-averaged expected directions of the geomagnetic field²⁵⁻²⁷ yields no evidence for tilting of the footwall. We conclude that the South Mountains master detachment fault was active as a low-angle extensional structure, with a dip of ~10°.

The principal structure of the South Mountains MCC (Fig. 1) is a subhorizontal, ductile to brittle, high-strain zone which accommodated large-magnitude extensional deformation. Initial mylonitic ductile deformation and later brittle detachment faulting were partitioned along a relatively thin zone (~100 m thick). Both phases of deformation are kinematically coordinated and indicate shear such that the top moves east-northeast. Ductile and brittle deformation occurred at relatively shallow crustal levels (within 5 to 10 km of the surface), and was promoted by magmatic heat and expulsion of magmatic fluids from a suite of Miocene synkinematic intrusive rocks^{22,23}. This synkinematic, footwall intrusive suite constitutes the northeastern part of South Mountains and includes (from oldest to youngest) South Mountains granodiorite (Tsm), Telegraph Pass granite (Ttp), felsic dykes (Tfd) and microdiorite dykes (Tmd)^{18,19}.

Palaeomagnetic data from undeformed portions of South Mountains footwall rocks provide a passive linear marker to

Received 17 August; accepted 21 November 1992.

1. Haddon, R. C. et al. *Nature* **350**, 320 (1991).
2. Xiang, X.-D. et al. *Science* **256**, 1190 (1992).
3. Ogata, H. et al. *Jpn J. appl. Phys.* **31**, L166 (1992).
4. Aslamasov, L. G. & Larkin, A. I. *Phys. Lett.* **A26**, 238 (1968).
5. Skocpal, W. J. & Tinkham, M. *Rep. Prog. Phys.* **38**, 1049 (1975).

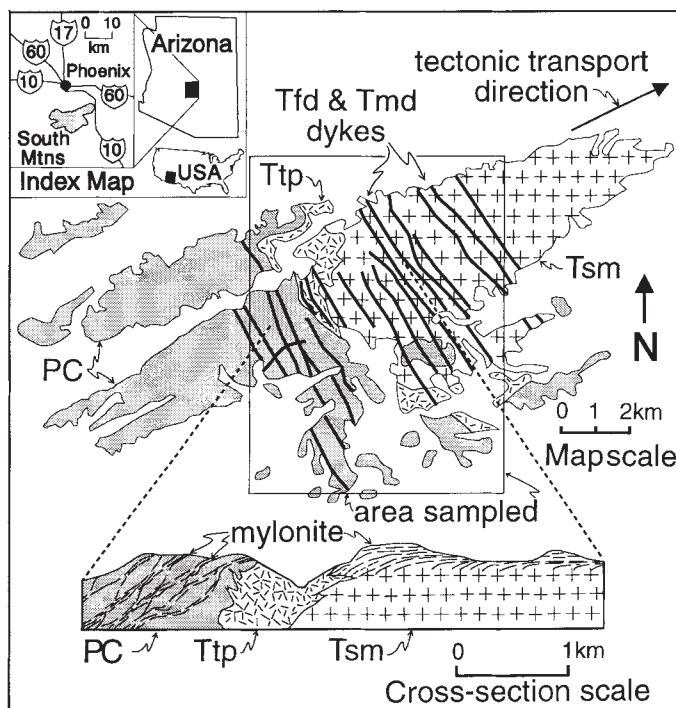


FIG. 1 Location map, generalized geological map and interpretive cross-section of the South Mountains metamorphic core complex, Arizona (PC, Precambrian rocks; Miocene age rocks: Tsm, South Mountains granodiorite; Ttp, Telegraph Pass granite, Tfd, felsic dykes; Tmd, microdiorite dykes). Palaeomagnetic data presented here are from 38 sites located in the central portion of South Mountains.

assess footwall tilting. Here we report palaeomagnetic data from a total of 38 sites, representing all phases of the South Mountains intrusive suite and Precambrian wallrocks (Figs 2, 3 and 4). The magnitude of footwall tilting is determined by comparing observed data (either primary thermoremanent magnetizations, TRMs, or high-temperature thermochemical magnetizations, TCRMs, acquired early in the cooling history of the intrusions) with expected, time-averaged directions of similar age²⁵⁻²⁷. Several palaeomagnetic studies in the Basin and Range province have focused on vertical axis rotation and on tilting of near-surface hanging wall rocks above master detachment faults²⁸⁻³². Here, in contrast, we focus on horizontal axis tilting of mid-crustal footwall rocks beneath a master detachment fault.

Demagnetization behaviour indicates that those well-grouped magnetizations (with high unblocking temperatures and high to moderate coercivities) that are found in South Mountains granodiorite and microdiorite dykes reside mainly in magnetite (Fig. 2). Those characteristic of the Telegraph Pass granite and

felsic dykes reside in both magnetite and hematite (Fig. 2). South Mountains granodiorite is of normal polarity, whereas the Telegraph Pass granite, microdiorite and felsic dykes are of both normal and reverse polarity (Fig. 3). The presence of normal or reverse polarity magnetizations is not a function of the magnetic phase dominating the remanence. Contact tests, involving dykes and host rocks, as well as between Telegraph Pass granite and Precambrian host rocks, have been attempted. In some of the tests involving dykes, the host rock has a polarity opposite to that of the dykes (Fig. 3). In these cases, rocks immediately adjacent to the dykes are partly remagnetized in the direction of the dyke. Some tests are inconclusive in that host rock and dyke give the same polarity and statistically indistinguishable direction of magnetization. Precambrian rocks adjacent to the Telegraph Pass granite give well-grouped magnetizations identical in direction to that of the granite (reverse polarity). Precambrian rocks at least 100 m away from the Telegraph Pass granite contact are of normal polarity (Fig. 3, PC host no. 4). Contact relationships 1 and 2 of Fig. 3 suggest that magnetizations carried by magnetite in South Mountains granodiorite and microdiorite dykes are primary TRMs. Contact relationships 3 and 4 of Fig. 3 suggest that hematite in the Telegraph Pass granite and felsic dykes is of high-temperature TCRM origin, with remanence acquisition occurring early in the cooling history of these intrusions. This hematite is most likely to have been formed during synkinematic circulation of moderate temperature ($\sim 600^\circ\text{C}$) magmatic fluids²³. Laboratory data on the unblocking temperature of the Telegraph Pass granite and felsic dykes do not allow us to define the temperature range over which magnetizations carried by hematite were blocked. For primary magnetizations in Telegraph Pass granite and felsic dykes carried dominantly by magnetite, most of the remanence was blocked between $\sim 400^\circ\text{C}$ and 580°C .

The thermal history of the South Mountains Granodiorite, provided by isotopic (U-Pb zircon, $^{40}\text{Ar}/^{39}\text{Ar}$ hornblende, K-Ar biotite) and fission-track (apatite) age determinations, indicates rapid cooling of $\sim 260^\circ\text{C Myr}^{-1}$ (from 800 to 300°C) between 22 and 20 Myr ago, followed by slower cooling of $\sim 73^\circ\text{C Myr}^{-1}$ (from 300 to 100°C) between 20 and 17 Myr (ref. 24). Mylonitization of footwall granitic intrusions (Tsm and Ttp) began shortly after crystallization (22 Myr)^{18,19,24}. Rapid cooling between 22 and 20 Myr reflects both uplift during ductile phases of extensional deformation and ambient cooling of the granitic intrusions²⁴. We therefore interpret magnetizations of high unblocking temperature and high to moderate coercivity found in undeformed parts of the granitic intrusions (Tsm and Ttp) as having been acquired just before or during mylonitization of structurally higher rocks. Felsic dykes that cut strongly mylonitized portions of the granitic intrusions (Tsm and Ttp) are only weakly mylonitized. Therefore, magnetizations of high unblocking temperature and high to moderate coercivity found in undeformed parts of the felsic dykes are interpreted to have been acquired during the final phases of ductile deformation of structurally higher rocks. Slower cooling below 300°C , between

FIG. 2 Representative modified progressive demagnetization diagrams (Roy plots) for South Mountains granodiorite (Tsm; NRM to 600°C), Telegraph Pass granite (Ttp; NRM to 670°C), felsic dyke (Tfd; NRM to 560°C), and microdiorite dyke (Tmd; NRM to 140 mT). Solid dots are vector end points that intersect the horizontal plane (declination), open squares represent vector end points that intersect the true vertical plane (horizontal against vertical, inclination).

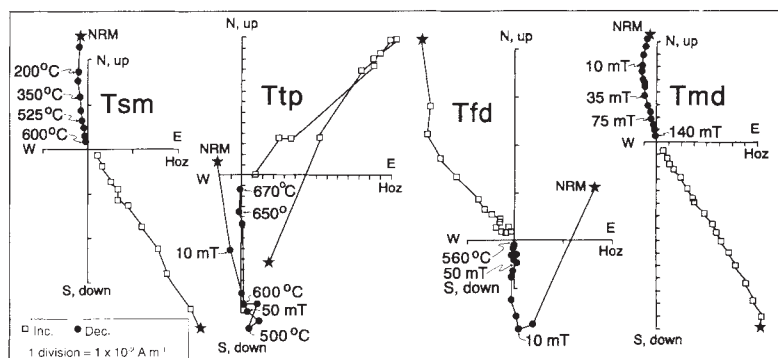
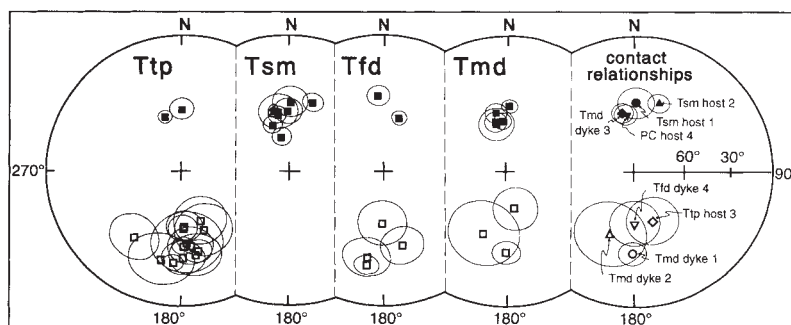


FIG. 3 Equal-area projections of site means with α_{95} confidence cones for Telegraph Pass granite (Ttp, 15 sites; this includes two sites of reverse-polarity Precambrian host rocks found within 10 m of the contact, and one site of reverse-polarity mafic xenoliths found within the Telegraph Pass granite), South Mountains granodiorite (Tsm, eight sites), felsic dykes (Tfd, six sites) and microdiorite dykes (Tmd, eight sites). Solid symbols represent lower-hemisphere projections and open symbols represent upper-hemisphere projections. The following contact relationships are also illustrated: South Mountains granodiorite sites, Tsm host 1 and Tsm host 2, intruded by microdiorite dykes, Tmd dyke 1 and Tmd dyke 2, respectively; Telegraph Pass granite site, Ttp host 3, intruded by microdiorite dyke, Tmd dyke 3; and remagnetized Precambrian host rocks, PC host 4 (this site is not shown on other equal-area projections of this figure), intruded by felsic dyke, Tfd dyke 4.



20 and 17 Myr, is correlated with brittle phases of extension²⁴. Microdiorite dykes cross-cut the granitic intrusions (Tsm and Ttp), felsic dykes and mylonites. Microdiorite dykes that intersect the South Mountains detachment fault are highly brecciated and must therefore postdate ductile deformation (22 to 20 Myr) and be synkinematic with respect to brittle deformation (20 to 17 Myr)^{18,19}. Magnetizations of high unblocking temperature and high to moderate coercivity acquired by microdiorite dykes are interpreted to have been acquired synchronously with brittle phases of detachment faulting. Fission-track age determinations on apatites from undeformed granodiorite indicate that the South Mountains MCC has resided at temperatures of less than 100 °C since 17 Myr (termination of major deformation)²⁴.

In summary, the demagnetization behaviour, contact relationships and thermal history of South Mountains footwall rocks indicate acquisition (early in the cooling history of the intrusions) and preservation of primary TRMs and high-temperature TCRMs. We interpret the moderate dispersion and dual polarity (except for Tsm) of palaeomagnetic data to indicate that cooling and remanence acquisition occurred over a period long enough to average palaeosecular variation. The data do not support the idea that remanent magnetization was acquired during a post-extension thermal or low-temperature chemical event at shallow crustal levels.

Group and grand mean palaeomagnetic directions for the South Mountains intrusive suite are statistically (95% confidence) indistinguishable from Miocene reference directions (Fig. 4). Assuming that magnetizations in the South Mountains granodiorite, Telegraph Pass granite and felsic dykes were

acquired before and during ductile extensional deformation, we interpret these data as demonstrating that the South Mountains footwall has not been significantly tilted after mylonitic deformation. Similarly, assuming that magnetizations in the microdiorite dykes were acquired during initial stages of brittle deformation, we interpret these data as demonstrating that the South Mountains footwall has not been significantly tilted after the brittle stages of deformation. The dip of this master detachment fault at the time of active deformation is then assumed to coincide with its current dip of ~10° northeastwards. Large-magnitude extensional deformation must therefore have been accommodated along a low-angle detachment (normal) fault during both ductile and brittle phases of deformation.

Development of low-angle detachment structures is favoured in a rheologically layered continental crust¹⁰⁻¹⁶. Detachment structures initiate at moderate dips in brittle upper crust and flatten at mid-crustal levels (~10–15 km depth), because stress axes rotate as a result of ductile (perhaps fluid-like¹⁶) flow of a weak middle crust (beneath the brittle to ductile transition zone)¹⁰⁻¹⁶. In the case of South Mountains, synkinematic intrusions and associated high magmatic fluid pressures²³ may have raised the brittle to ductile transition zone to between 5 and 10 km depth. This would allow flattening of the dip angle of a detachment structure at higher crustal levels. Elsewhere, the dip of this detachment structure may have flattened at deeper crustal levels, resulting in an along-strike warped geometry, with only the structurally highest level (the South Mountains) currently exposed at the surface. Variations in dip of master detachment fault structures, ranging from low angle (~10° for South

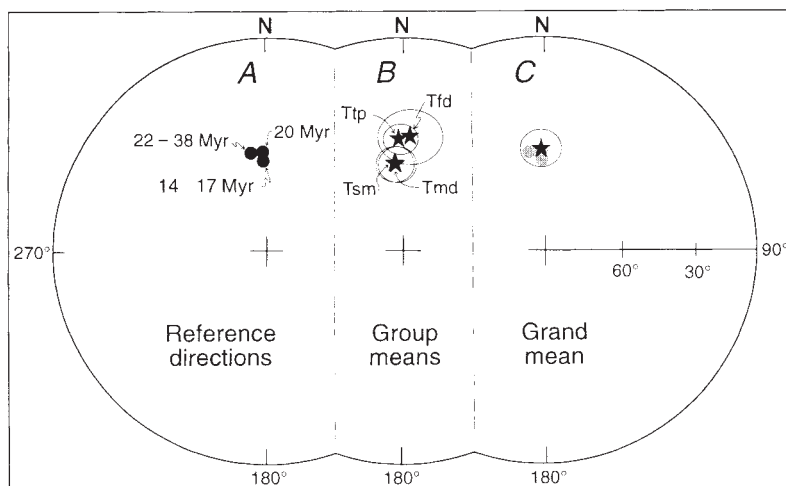


FIG. 4 a, Equal-area projection of Miocene reference directions for South Mountains, Arizona (solid dots, lower-hemisphere projections), calculated from palaeomagnetic pole data of Diehl *et al.*²⁵ (22 to 38 Myr), Irving and Irving²⁶ (20 Myr), and Mankinen *et al.*²⁷ (14 to 17 Myr). b, Equal-area projection of group-mean data (stars, lower-hemisphere projections; reverse-polarity data have been inverted to normal polarity) with α_{95} confidence cones for South Mountains granodiorite (Tsm), Telegraph Pass granite (Ttp), felsic dykes (Tfd), and microdiorite dykes (Tmd). c, Equal-area projection of grand mean (star, lower-hemisphere projection) with α_{95} confidence cone and the three Miocene reference directions (grey dots). Grand mean is derived from the four group-mean data points of b. The grand mean direction is statistically indistinguishable from the three Miocene reference directions, suggesting that there has been no footwall tilting. Therefore, the current dip of the high strain zone, ~10° northeastwards, is assumed to represent the dip of this structure during deformation.

Mountains and Sevier Desert detachments) to high angle, may then be controlled by lateral variations in thermal and mechanical properties of a layered crust¹⁰⁻¹⁶. □

Received 11 May; accepted 10 November 1992.

1. Miller, E. L., Gans, P. B. & Garling, J. *Tectonics* **2**, 239-263 (1983).
2. Davis, G. H. *Geology* **11**, 342-347 (1983).
3. Jackson, J. A. *Continental Extensional Tectonics*, *Geol. Soc. spec. Publs* **28**, 3-17 (1987).
4. Jackson, J. A. & White, N. J. *J. struct. Geol.* **11**, 15-36 (1989).
5. King, G. & Ellis, M. *Nature* **348**, 689-693 (1990).
6. Miller, M. G. *Geology* **19**, 372-375 (1991).
7. Wernicke, B. & Axen, G. J. *Geology* **16**, 848-851 (1988).
8. Hamilton, W. B. *U.S. Geol. Surv. Bull.* **1790**, 51-85 (1988).
9. Buck, W. R. *Tectonics* **7**, 959-973 (1988).
10. Davis, G. A. & Lister, G. S. *Geol. Soc. Am. spec. Pap.* **218**, 133-159 (1988).
11. Lister, G. S. & Davis, G. A. *J. struct. Geol.* **11**, 65-94 (1989).
12. Yin, A. *Tectonics* **8**, 469-482 (1989).
13. Reynolds, S. J. & Lister, G. S. *Geology* **18**, 216-219 (1990).
14. Melosh, H. J. *Nature* **343**, 331-335 (1990).
15. Forsyth, D. W. *Geology* **20**, 27-30 (1992).
16. Wernicke, B. *The Cordilleran Orogen: Coterminous United States, The Geology of North America* G3, 553-572 (Geological Society of America, Boulder, Colorado, 1992).
17. Spencer, J. E. *Geology* **12**, 95-98 (1984).
18. Reynolds, S. J. & Rehrig, W. A. *Geol. Soc. Am. Mem.* **153**, 159-175 (1980).
19. Reynolds, S. J. *Arizona Bur. Geol. Min. Tech. Geol. Sur. Branch Bull.* **195** (1985).
20. Reynolds, S. J., Shafiqullah, M., Damon, P. E. & DeWitt, E. *Geology* **14**, 283-286 (1986).
21. Davis, G. A., Lister, G. S. & Reynolds, S. J. *Geology* **14**, 7-10 (1986).
22. Reynolds, S. J., Richard, S. M., Haxel, G. B., Tosdal, R. M. & Laubach, S. E. *Metamorphism and Crustal Evolution, Western Conterminous United States, Rubey 7*, 466-501 (Prentice-Hall, Englewood Cliffs, 1988).
23. Smith, B. M., Reynolds, S. J., Day, H. W. & Bodnar, R. J. *Geol. Soc. Am. Bull.* **103**, 559-569 (1991).
24. Fitzgerald, P. G., Reynolds, S. J., Stump, E., Foster, D. A. & Gleadow, A. J. W. *Tectonics* (submitted).
25. Diehl, J. F., McClannahan, K. M. & Bornhorst, T. J. *J. geophys. Res.* **93**, 4869-4879 (1988).
26. Irving, E. & Irving, G. A. *Geophys. Surv.* **5**, 141-188 (1982).
27. Markinen, E. A., Larson, E. E., Gromme, C. S., Prevot, M. & Coe, R. S. *J. geophys. Res.* **92**, 8057-8076 (1987).
28. Hagstrum, J. T. & Gans, P. B. *J. geophys. Res.* **94**, 1827-1847 (1989).
29. Hudson, M. R. & Geissman, J. W. *J. geophys. Res.* **96**, 3979-4006 (1991).
30. Jancke, S. U., Geissman, J. W. & Bruhn, R. L. *Tectonics* **10**, 403-432 (1991).
31. Faulds, J. E., Geissman, J. W. & Shafiqullah, M. *Tectonics* **11**, 204-227 (1992).
32. Doughty, T. P. & Sheriff, S. D. *Tectonics* **11**, 663-671 (1992).

ACKNOWLEDGEMENTS. We thank S. B. Keith and S. Ruff of MagmaChem Exploration for their hospitality, and D. Bahar, R. F. Butler, M. J. Grubensky and B. Horton for field assistance. This work was funded by Sigma Xi, the Student Research Allocation Committee at the University of New Mexico and the NSF.

Patterns of fine root mortality in two sugar maple forests

Ronald L. Hendrick & Kurt S. Pregitzer*

Department of Forestry, Michigan State University, E. Lansing, Michigan 48824, USA

MUCH of the carbon assimilated by plants is allocated to fine root production¹⁻⁵, and the amount of carbon and nutrients subsequently returned to the soil from fine root turnover equals or surpasses that returned through leaf litter in many forests⁶⁻⁹. Unfortunately, limitations in traditional methods of studying roots have prevented us from thoroughly understanding the dynamic nature of fine root mortality in most forests, and better measurements of fine root longevity are needed to quantify and model more accurately ecosystem carbon and nutrient budgets⁸⁻¹¹. We used minirhizotrons^{12,13} to follow the mortality of contemporaneous fine root cohorts in two sugar maple (*Acer saccharum* Marsh.) forests located 80 km apart (north-south) during 1989 and 1990. We report here that roots in the northern forest consistently lived the longest, principally owing to greater rates of mortality early in the life of roots at the southern site. Differences in site factors suggest that warmer soil temperatures seem to be associated with the more rapid death of roots at the southern site.

During the 1989 and 1990 growing seasons, we used a micro-video camera^{12,13} to record monthly images of fine root cohorts growing along the exterior surfaces of 12 minirhizotrons (clear plastic tubes) located in each forest. Fine root cohorts were comprised of all living roots ≤ 1.5 mm in diameter (pooled across

the 12 minirhizotrons¹⁴) that were produced during the intervals between sampling dates. We used a microcomputer software program (ROOTS^{12,15}) to identify and measure the length of individual roots within contemporaneous cohorts at each site, and subsequently to follow their fate by repeatedly measuring the same roots at each sampling date.

Because there are no appreciable differences among the sites in soil organic matter content, exchangeable Ca, Mg, Al or K¹⁶, N mineralization, growing season soil moisture content, canopy leaf area, litterfall mass or N^{17,18}, we hypothesized that fine root length would be lost at similar rates in both forests. Instead, we found large differences in fine root lifespans (Fig. 1). Analyses of survival distributions of contemporaneous cohorts showed that roots at the northern site consistently lived longer on average than roots born during the same periods at the southern site (Gehan-Mantel¹⁹ test, $\alpha = 0.05$). The only exceptions were in July (no site difference) and September (greater longevity at the southern site) of 1989. The longer lifespan of roots in the northern forest was due to significantly lower first-season mortality rates (Fig. 2). New roots were lost 64% faster at the southern forest (0.41 versus 0.25% per day at the northern forest, $P < 0.01$), but overwinter and second-year mortality rates were not significantly different ($P > 0.5$; 0.14 versus 0.12% per day at the northern forest).

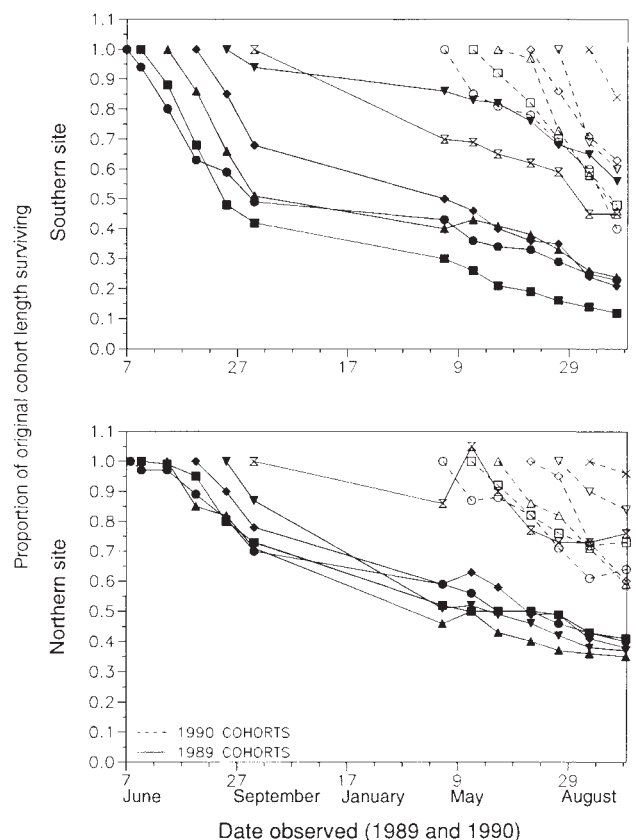


FIG. 1 Root length survival curves for root cohorts during 1989 and 1990. Survival times were significantly greater at the northern site for all but the July (no difference) and October (southern site greater) 1989 cohorts. Cohorts were comprised of all new roots present along a 1.8-cm-wide transect on the minirhizotron surface that were produced between successive sample dates in the upper 30 cm of the soil. The data were pooled from the 12 minirhizotrons in each forest. Cohort sizes ranged from 26 to 245 roots, with a mean and standard deviation of 87 and 50, respectively. Each root was identified and numbered by the ROOTS image analysis program we developed, and was remeasured and classified as live or dead at each sampling date. Significant differences in length survival were tested using a Gehan-Mantel test for the censored data¹⁹ at an α of 0.05.

* To whom correspondence should be addressed.

The good fits of the first growing season mortality regressions (Fig. 2) were due to very similar patterns of mortality for all cohorts at each site (Fig. 1). These results were unexpected, because several studies have shown seasonal variation in forest ecosystem fine root biomass turnover^{2,3,20-22}, which is often believed to reflect variable root mortality rates. But the dynamics of root length and root biomass production and mortality are similar within both forests (R.L.H. and K.S.P., manuscript in preparation) and our results suggest that temporal differences in biomass 'turnover' may be due to temporal variation in root production, not in root mortality, in ecosystems like those we studied.

Low overwinter fine root mortality, relative to growing season mortality, might be due to low maintenance respiration rates under cold soil temperatures²³⁻²⁵, but it is unclear why roots produced in the 1989 growing season should continue to die at a slower rate during the 1990 growing season. Survival of these roots may have been enhanced by greater lignification and/or suberization of cortical or epidermal cells, or perhaps they were located in resource-rich microsites. Low branch orders of roots can also live longer than higher order roots²⁶. Our study was not designed to address these possibilities, but they deserve further investigation.

Both forests we studied are similar in structure, composition and site quality (Table 1 and refs 14-16). Thus, site differences in these factors cannot be invoked as obvious explanations^{1,3,9} for the site differences in mortality rates and survival distributions. But it is possible that warmer soil temperatures contribute in part to greater fine root mortality rates in the southern forest (Table 1). Late-spring and summer soil temperatures at 15-cm depth were 2-4 °C warmer at the southern site in both 1989 and 1990 (Fig. 3), and soil temperatures were significantly warmer at the southern site from late April to mid-October during both years (paired *t*-test; *P* < 0.001) by an average of about 1.5 °C. Soil temperature was not controlled in our study, and so the evidence linking temperature and root mortality is circumstantial. But previous studies have demonstrated increased root respiration rates and faster root turnover associated with higher soil temperature²³⁻²⁵, and we believe that higher soil temperatures at the southern forest act in part to accelerate maintenance costs and mortality rates throughout the entire system of

TABLE 1 Stand, soil and climatic characteristics of the study sites

	Southern forest	Northern forest
General stand characteristics		
Latitude	43°40'	44°23'
Longitude	86°09'	85°50'
Above-ground biomass (Mg ha ⁻¹)	234	275
Overstory age (yr)	78 (3)	74 (6)
Basal area (m ² ha ⁻¹)	30 (1.2)	30 (1.6)
Sugar maple basal area (%)	75 (8)	83 (3)
Canopy height (m)	24 (1)	28 (1)
Leaf area index (m ² m ⁻²)		
1989	7.9 (0.2)	7.1 (0.5)
1990	7.7 (0.2)	7.8 (0.3)
Mean annual air temperature (°C)	7.6	5.8
Mean annual precipitation (mm)	850	810
Mean spring-fall soil temperature (°C at 15 cm)		
26/4/89-17/10/89	14.2 (4.3)	12.9 (3.5)
30/4/90-16/10/90	14.6 (3.3)	12.9 (3.2)
Soil variables (A + E Horizons)		
pH (1:1 water)	4.66 (0.046)	4.66 (0.27)
Bulk density (g cm ⁻³)	1.27 (0.07)	1.27 (0.04)
Nitrogen (kg ha ⁻¹)	1,408 (100)	1,125 (398)
N mineralization (mg kg ⁻¹ yr ⁻¹)	134.4	100.8
Phosphorus (kg ha ⁻¹)	156 (18)	191 (58)
Organic matter (Mg ha ⁻¹)	41.7 (3.3)	45.7 (13.6)

Values are averages across plots, with s.e.m. in parentheses where shown.

fine, absorbing roots. Soil organisms, which would be more active in the warmer soil, can also shorten root longevity²⁷.

The relationship between soil temperature and root longevity needs to be more firmly established, and rates of fine root mortality may not be as predictable in other ecosystems as they are in the forests we studied. Nonetheless, our findings highlight the need for a better understanding of the processes controlling fine root inputs of carbon and nutrients to the soil. For example, below-ground resource allocation is driven by nitrogen or water availability^{28,29} in many forest ecosystem models. In others,

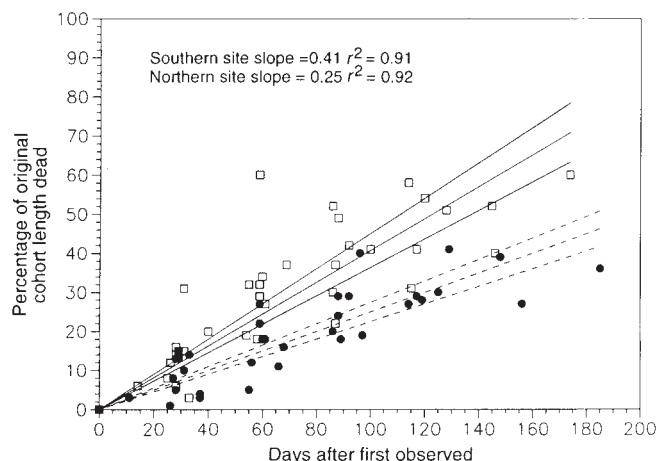


FIG. 2 Regressions of length mortality data against the number of days a cohort had been under observation during its first growing season, with the line forced through the origin. The slope (mortality rate) of the southern site was significantly greater than the northern site. There were no site differences in over-winter and second-year mortality rates (regressions not shown). Data were pooled across all 1989 and 1990 cohorts. Differences among the two forests were tested by comparing regression slopes at an α of 0.05. 95% confidence intervals shown.

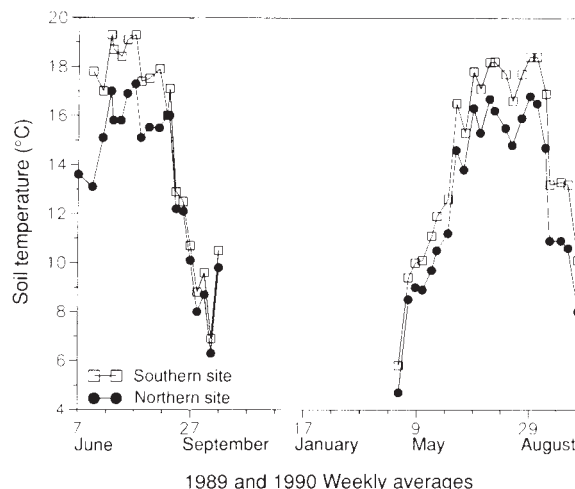


FIG. 3 Weekly averages of soil temperatures at 15-cm depth in both forests. Paired *t*-tests of site differences in weekly averages (derived from the mean of three temperature sensors in each forest) showed that temperatures from 26/4/89 to 17/10/89 and 30/4/90 to 16/10/90 were significantly higher in the southern forest (*P* < 0.001). Data were collected every 30 min with an Omnidata monitoring system (Omnidata International, Logan, Utah, USA).

carbon and nutrient budgets allow for greater C costs through increased root respiration in warmer soils^{10,29,30} associated with seasonal or climatic increases in temperature. But none of these models or any others that we are aware of make allowance for possible differences in root turnover rates within an ecosystem type, nor are root turnover rates intrinsically linked to soil temperature. We do not wish to imply that current models are flawed in their construction. Instead, they reflect how little is known of the demographic processes (such as root mortality) that control below-ground carbon and nutrient cycling, and the relationship between root demography and the soil environment. Given the importance of fine roots in terrestrial carbon and nutrient budgets¹⁻¹⁰, the need for more and better data on fine root dynamics is apparent. Even if soil temperature is generally found to have a minimal influence on root mortality, the fact that rates of fine root turnover can vary substantially among virtually identical ecosystems needs to be reconciled with current knowledge and models of below-ground processes that do not account for this phenomenon. □

3. Keyes, M. R. & Grier, C. C. *Can. J. For. Res.* **11**, 599-605 (1981).
4. Harris, W. F., Kinerson, R. S. & Edwards, N. T. *Pedobiologia* **17**, 369-381 (1977).
5. Fogel, R. in *Ecological Interactions in Soil* (eds. Fitter, A. H., Atkinson, D., Read, D. J. & Usher, M. B.) 23-36 (Blackwell Scientific, Oxford, 1985).
6. Cox, T. L., Harris, W. F., Ausmus, B. S. & Edwards, N. T. *Pedobiologia* **18**, 264-271 (1978).
7. Joslin, J. D. & Henderson, G. S. *For. Sci.* **33**, 330-346 (1987).
8. Raich, J. W. & Nadelhoffer, K. J. *Ecology* **70**, 1346-1354 (1989).
9. Ewel, K. C. & Gholz, H. L. *For. Sci.* **37**, 397-438 (1991).
10. Landsberg, J. J., Kaufmann, M. R., Binkley, D., Isebrands, J. & Jarvis, P. G. *Tree Physiol.* **9**, 1-15 (1991).
11. Hendrick, R. L. & Pregitzer, K. S. *Ecology* **73**, 1094-1104 (1992).
12. Upchurch, D. R. & Ritchie, J. T. *Agron. J.* **75**, 1009-1015 (1983).
13. Hendrick, R. L. & K. S. Pregitzer. *Pl. Soil* **143**, 283-288 (1992).
14. Atkinson, D. *Trends Ecol. Evol.* **7**, 173-174 (1992).
15. MacDonald, N. W., Burton, A. J., Jurgensen, M. F., McLaughlin, J. W. & Morz, G. D. *Soil Sci. Soc. Am. J.* **55**, 1709-1715 (1991).
16. Burton, A. J., Pregitzer, K. S. & Reed, D. D. *For. Sci.* **37**, 1041-1059 (1991).
17. Pregitzer, K. S. & Burton, A. J. *Can. J. For. Res.* **21**, 1148-1153 (1991).
18. Pyke, D. A. & Thompson, J. N. *Ecology* **67**, 240-245 (1986).
19. Fogel, R. *Pl. Soil* **71**, 75-85 (1983).
20. McClaugherty, C. A., Aber, J. A. & Melillo, J. M. *Ecology* **63**, 1481-1490 (1982).
21. Persson, H. *Vegetatio* **41**, 101-109 (1979).
22. Amthor, J. S. *Pl. Cell Environ.* **7**, 561-569 (1984).
23. Lawrence, W. T. & Oechel, W. C. *Can. J. For. Res.* **13**, 840-849 (1983).
24. Marshall, J. P. & Waring, R. H. *Can. J. For. Res.* **15**, 791-800 (1985).
25. Atkinson, D. in *Ecological Interactions in Soil* (eds Fitter, A. H., Atkinson, D., Read, D. J. & Usher, M. B.) 43-65 (Blackwell Scientific, Oxford, 1985).
26. Head, G. C. in *Shedding of Roots* (ed. T. T. Kozlowski) 237-293 (Academic, New York, 1973).
27. Aber, J. D., Melillo, J. M., Nadelhoffer, K. J., Pastor, J. & Boone, R. D. *Ecol. Appl.* **1**, 303-315 (1991).
28. Running, S. W. & Gower, S. T. *Tree Physiol.* **9**, 147-160 (1991).
29. Bonan, G. B. *J. geophys. Res.* **96**, 7301-7312 (1991).

Received 11 May; accepted 20 October 1992.

1. Gower, S. T., Vogt, K. A. & Grier, C. C. *Ecol. Monogr.* **62**, 43-65 (1992).
2. Grier, C. C., Vogt, K. A., Keyes, M. R. & Edmonds, R. L. *Can. J. For. Res.* **11**, 155-167 (1981).

ACKNOWLEDGEMENTS. We thank D. Eisenstat, R. Fogel, T. Gower, J. Marshall, B. Palik and C. Ramm for comments on the manuscript. This research was funded by the USDA Forest Service Northern Stations Global Change Program, the MERRA and the Michigan Agricultural Experiment Station.

Increase in C3 plant water-use efficiency and biomass over Glacial to present CO₂ concentrations

H. Wayne Polley*, Hyrum B. Johnson*,
Bruno D. Marino† & Herman S. Mayeux*

* Department of Agriculture, Agricultural Research Service,
808 East Blackland Road, Temple, Texas 76502, USA

† Department of Earth and Planetary Sciences and Division of Applied
Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

ATMOSPHERIC CO₂ concentration was 160 to 200 $\mu\text{mol mol}^{-1}$ during the Last Glacial Maximum (LGM; about 18,000 years ago)¹, rose to about 275 $\mu\text{mol mol}^{-1}$ 10,000 years ago^{2,3}, and has increased to about 350 $\mu\text{mol mol}^{-1}$ since 1800 (ref. 4). Here we present data indicating that this increase in CO₂ has enhanced biospheric carbon fixation and altered species abundances by increasing the water-use efficiency of biomass production of C₃ plants, the bulk of the Earth's vegetation. We grew oats (*Avena sativa*), wild mustard (*Brassica kaber*) and wheat (*Triticum aestivum* cv. Seri M82 and Yaqui 54), all C₃ annuals, and selected C₄ grasses along daytime gradients of Glacial to present atmospheric CO₂ concentrations in a 38-m-long chamber. We calculated parameters related to leaf photosynthesis and water-use efficiency from stable carbon isotope ratios (¹³C/¹²C) of whole leaves. Leaf water-use efficiency and above-ground biomass/plant of C₃ species increased linearly and nearly proportionally with increasing CO₂ concentrations. Direct effects of increasing CO₂ on plants must be considered when modelling the global carbon cycle and effects of climate change on vegetation.

The shoots of plants grown in the 38-m chamber were enclosed by a transparent cover through which air was moved in one direction. Photosynthesis by enclosed plants progressively depleted the CO₂ concentration ([CO₂]) and increased the ¹³C/¹²C of air (B.D.M. *et al.*, manuscript in preparation) as it moved the 38 m from the air intake to outlet of the chamber. The $\delta^{13}\text{C}$ values (see Fig. 1 legend) of leaves of C₃ species and of concurrently grown C₄ grasses, grassbur (*Cenchrus incertus*), crabgrass (*Digitaria ciliaris*) and Gaspé yellow flint maize (*Zea mays*),

were linearly correlated with [CO₂] (Fig. 1). The difference between the $\delta^{13}\text{C}$ of atmospheric CO₂ and leaf carbon of maize (and certain other C₄ species) is conservative across environments⁵ and did not vary significantly with [CO₂] (B.D.M. *et al.*, manuscript in preparation), enabling us to use the $\delta^{13}\text{C}$ of C₄ leaves as a proxy for the $\delta^{13}\text{C}$ of chamber air⁶.

Leaf intercellular [CO₂] (*c_i*), calculated from leaf $\delta^{13}\text{C}$ values, increased linearly and nearly proportionally (by the same ratio) with higher ambient [CO₂] (*c_a*) in each C₃ species. As a result, intercellular [CO₂] was a constant but different fraction of *c_a* in oats (*c_i/c_a* = 0.70) and mustard (*c_i/c_a* = 0.65) grown at mean [CO₂] from 330 to 160 $\mu\text{mol mol}^{-1}$ (Fig. 2). The *c_i/c_a* of wheat cultivars increased only slightly from 0.63 and 0.66 at 225 $\mu\text{mol mol}^{-1}$ to 0.66 and 0.69 at 350 $\mu\text{mol mol}^{-1}$ for Yaqui 54 and Seri M82, respectively. A similar result has been reported⁷ in young wheat plants exposed to CO₂ partial pressures between 120 and 500 μbar . Others have demonstrated that coordination of stomatal and mesophyll functions minimizes variation in *c_i/c_a* to similar values in C₃ species⁸, but this is the first demonstration of such for C₃ plants grown over a [CO₂] range characteristic of Last Glacial Maximum (LGM)-to-present atmospheric concentrations. Intrinsic water-use efficiency, defined as the ratio of leaf photosynthesis or net assimilation (*A*) to stomatal conductance to water vapour (*g*), increased by the same (oats and mustard) or nearly the same (wheat) relative amount as did [CO₂], a consequence of conservative *c_i/c_a* in these C₃ species (Fig. 3).

Leaf assimilation should have increased substantially at the higher *c_i* that accompanied increasing [CO₂], unless photosynthetic capacity (the relationship of *A* to *c_i*) declined. The positive linear relationship of *A* to *c_i* at subambient [CO₂] that is typical of C₃ species did not differ between oat plants grown at extremes of the [CO₂] gradient⁹. We estimate from that relationship that with a constant *c_i/c_a* of 0.70, net assimilation of oat leaves would have increased ~40% with the 75 $\mu\text{mol mol}^{-1}$ rise in [CO₂] since 1800. That the increase in leaf *c_i* was correlated with greater plant carbon gain is evidenced by the positive linear relationships between *c_i* and aboveground biomass per plant of all C₃s studied (Fig. 4).

Climate¹⁰, and particularly site water balance¹¹, largely control the structure and productivity of vegetation. This control is determined in part by plant water-use efficiency (WUE), biomass produced per unit transpiration. Our results imply that WUE of C₃ plants may have increased by 27% over the past 200 years

and ~100% since the LGM. Water-use efficiency calculated from $\delta^{13}\text{C}$ is often highly correlated with the ratio of biomass to transpiration of C3 plants¹², but the correlation is influenced by leaf-to-air vapour pressure differences. Little of the CO_2 -induced increase in potential WUE might be realized if higher A/g resulted entirely from stomatal closure. Then, transpirational cooling of leaves would decline and leaf temperature and transpiration rate per unit g would rise¹³. Any decline in water loss might also be partly offset by the resulting increase in atmospheric water vapour pressure deficit. But the c_i and biomass/plant of C3 annuals increased by the same ratio as $[\text{CO}_2]$, indicating that much of the CO_2 -induced increase in A/g (and WUE) was realized as higher A and ultimately as greater plant biomass (Fig. 4). This increase in plant A/g from the LGM to present must have extended the geographic ranges of some

species into areas where precipitation was formerly too low to support growth. We calculate from regression that A/g of oats and mustard would have increased 14.1 to 16.4 mmol CO_2 per mol H_2O , respectively, with the increase from 275 to 350 $\mu\text{mol mol}^{-1}$ CO_2 that has occurred since 1800. The increase is comparable to the mean 27.6% rise in A/g of 21.9 mmol mol^{-1} of C3 species from moist wash to drier slope habitats in the Sonoran desert, where the change in A/g was associated with a shift in species composition¹⁴.

Our results and those of others^{15,16}, if representative for C3 species, imply that the rise in $[\text{CO}_2]$ since the LGM greatly increased potential productivity of most of Earth's vegetation. Many believe, however, that higher $[\text{CO}_2]$ has a negligible effect

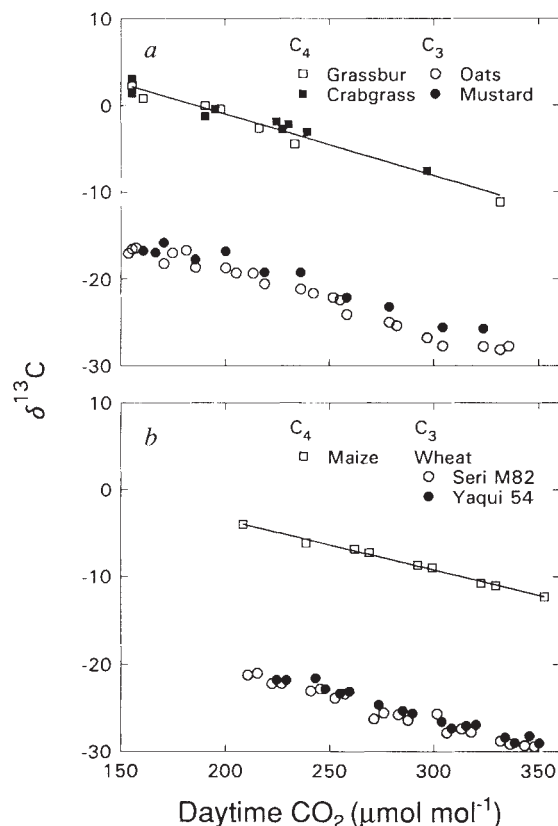


FIG. 1 The stable carbon isotope composition of upper leaves from C3 and C4 plants that were grown at daytime $[\text{CO}_2]$ from near 350 $\mu\text{mol mol}^{-1}$ to a, 150 (19 January to 3 May, 1989) or b, 200 $\mu\text{mol mol}^{-1}$ (12 February to 23 May, 1991). The C-isotope composition of whole leaves was determined by mass spectrometry (ISOMASS; VG Isogas) and expressed as $\delta^{13}\text{C}$, ‰ (parts per thousand) ^{13}C relative to a Pee Dee belemnite reference standard: $\delta^{13}\text{C} = [(^{13}\text{C}/^{12}\text{C})_{\text{sample}} / (^{13}\text{C}/^{12}\text{C})_{\text{standard}} - 1] \times 10^3$. Lines are linear regressions of $\delta^{13}\text{C}$ of C4 species on mean daytime $[\text{CO}_2]$ during growth: $\delta^{13}\text{C} = 13.30 - 0.07[\text{CO}_2]$, $r^2 = 0.96$, $P < 0.0001$ for grassbur and crabgrass, and $\delta^{13}\text{C} = 7.91 - 0.06[\text{CO}_2]$, $r^2 = 0.99$, $P < 0.0001$ for maize. Plants were grown from seed in a 38-m long chamber in a ventilated glasshouse²⁷. Soil water was restored weekly to field capacity. The chamber consisted of a 0.6-m-high polyethylene cover attached to the top of a 0.76-m-deep and 0.45-m-wide soil container. A desired $[\text{CO}_2]$ gradient was maintained in the chamber during daylight (9–10 h daily) by automatically varying the rate of unidirectional air flow through the cover in response to changes in photosynthetic CO_2 depletion by enclosed plants and sunlight intensity. Standard deviation of the minimum daytime $[\text{CO}_2]$ was less than 35 $\mu\text{mol mol}^{-1}$ on more than 80% of days. Drybulb and dewpoint temperatures of air were regulated at each 7.6 m along the chamber to track temporal variation in the glasshouse. The daytime drybulb temperature and vapour pressure deficit of air after regulation at each 7.6 m along the chamber were a mean 26°C ($N = 34$) and 1.94 kPa ($N = 33$) during the time sampled leaves of oats and mustard expanded and 20.6°C ($N = 24$) and 1.13 kPa ($N = 19$) during wheat growth.

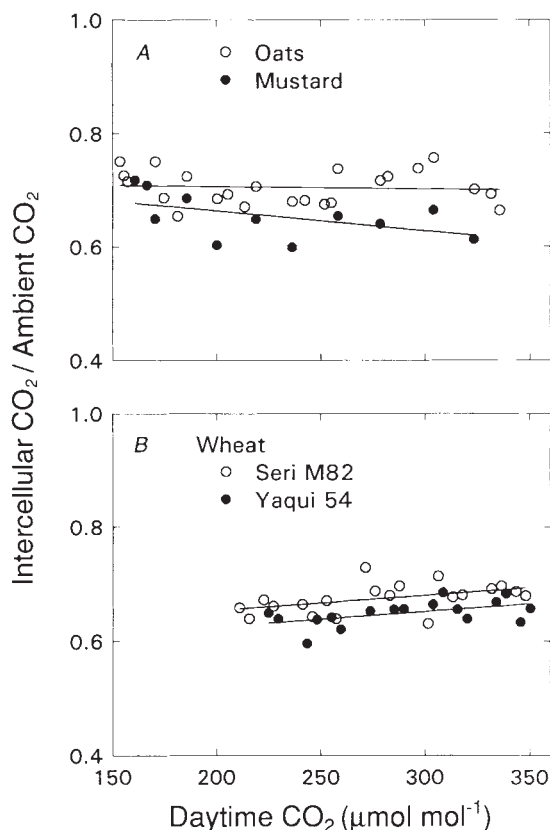


FIG. 2 The ratio of leaf intercellular $[\text{CO}_2]$ (c_i) to ambient $[\text{CO}_2]$ (c_a), calculated from the $\delta^{13}\text{C}$ of leaves, as a function of the daytime $[\text{CO}_2]$ at which C3 plants were grown. The slopes of regressions of c_i/c_a on c_a did not differ significantly from zero for (A) oats ($P = 0.70$) or mustard ($P = 0.11$); the c_i/c_a of B, the two wheat cultivars, was positively related to c_a (Seri M82, $c_i/c_a = 0.60 + (2.66 \times 10^{-4})c_a$, $r^2 = 0.22$, $P = 0.04$ and Yaqui 54, $c_i/c_a = 0.57 + (2.71 \times 10^{-4})c_a$, $r^2 = 0.27$, $P = 0.04$). The stable C isotope composition of plants ($\delta^{13}\text{C}$) reflects that of atmospheric CO_2 ($\delta^{13}\text{C}$) and fractionation during photosynthesis²⁸. For C3 plants, $\delta^{13}\text{C} = \delta^{13}\text{C}_a - a - (b - a)c_i/c_a$, where a is a fractionation resulting from the greater diffusivity of $^{12}\text{CO}_2$ than $^{13}\text{CO}_2$ across leaf stomata (4.4‰ (ref. 29); positive values indicate discrimination against ^{13}C or ^{12}C -enrichment) and b is discrimination by ribulose 1,5-bisphosphate carboxylase against ^{13}C in CO_2 fixation (29‰; ref. 30). Leaf c_i/c_a can be deduced from the above equation when $\delta^{13}\text{C}$ and $\delta^{13}\text{C}_a$ are known. We calculated $\delta^{13}\text{C}_a$ as a function of c_a in the 38-m chamber from the $\delta^{13}\text{C}$ of leaves from C4 plants that were grown with oats, mustard and wheat using a fractionation by C4 plants relative to air of 3.136‰. Isotope fractionation by maize grown across a 350–200 $\mu\text{mol mol}^{-1}$ CO_2 gradient in an additional experiment did not vary significantly with $[\text{CO}_2]$ ($\bar{x} = 3.136\%$, s.e. = 0.06%, $N = 9$; B.D.M. *et al.*, manuscript in preparation). Regressions of $\delta^{13}\text{C}$ on daytime $[\text{CO}_2]$ did not differ significantly for grassbur and crabgrass grown with oats and mustard (Fig. 1a; $F_{(2,12)} = 2.6$, $P < 0.10$) or for maize and crabgrass grown with wheat (data not shown; $F_{(2,29)} = 2.7$, $P > 0.10$). The small changes in drybulb and dewpoint temperatures between points of regulation at each 7.6 m along the chamber⁹ had no significant influence on the relationship of c_i/c_a of individual species or cultivars to c_a .

on global carbon fixation because factors other than CO_2 limit productivity of most unmanaged vegetation^{17,18}. Some recent research suggests that a step increase in $[\text{CO}_2]$ above the current $350 \mu\text{mol mol}^{-1}$ produces little or no short-term (≤ 3 years) increase in plant or ecosystem carbon storage where low temperatures or nutrient (particularly nitrogen) availability currently restrict plant growth^{19,20}. Extrapolation from these studies to the past is difficult, for only superambient $[\text{CO}_2]$ and a limited range of processes and temporal scales were considered. Plant water- and light-use efficiencies²¹ were lower at subambient $[\text{CO}_2]$, implying that sustainable biomass and plant nutrient requirements were also lower in the past. Species and genetic change, fixation of atmospheric N_2 , and nitrogen deposition may have facilitated plant response to CO_2 in the past when concentrations rose more slowly or with a greater relaxation time between change than today, but data are lacking.

Increased widths of annular rings of some trees²², global CO_2 models²³, calculations of carbon accumulation in temperate forests²⁴, and the increased amplitude of the annual cycle of atmospheric $[\text{CO}_2]$ ²⁵ in recent decades, all suggest that rising $[\text{CO}_2]$ has stimulated biospheric carbon fixation. Effects of $[\text{CO}_2]$, however, cannot readily be distinguished from those of human impact and concurrent climate change²⁴. Resolution of what fraction has been realized, if any, of the potentially great increase in plant productivity since the LGM demands that effects of $[\text{CO}_2]$ on the processes that influence plant growth be

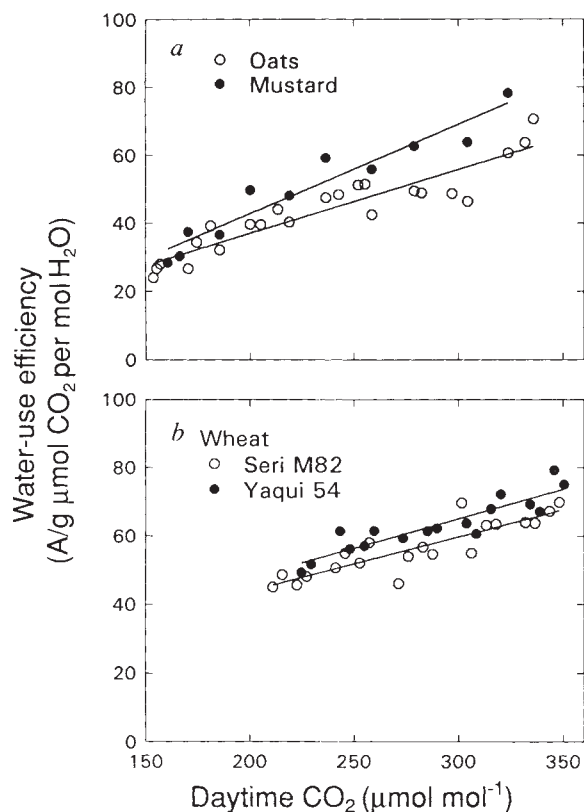


FIG. 3 Relationships between intrinsic water-use efficiencies of C3 plants (net assimilation (A)/stomatal conductance to water (g)) derived from the $\delta^{13}\text{C}$ of leaves and the daytime $[\text{CO}_2]$ (c_a) at which plants were grown. Linear regressions differed significantly ($P < 0.005$) between *a*, oats ($A/g = -0.70 + 0.19c_a$, $r^2 = 0.86$) and mustard ($A/g = -9.91 + 0.26c_a$, $r^2 = 0.92$), and *b*, two cultivars of wheat, Seri M82 ($A/g = 12.05 + 0.16c_a$, $r^2 = 0.77$) and Yaqui 54 ($A/g = 13.29 + 0.17c_a$, $r^2 = 0.82$). $P < 0.0001$ for each regression. Intrinsic water-use efficiency of C_3 leaves is directly proportional to c_a and negatively correlated with the ratio of leaf intercellular $[\text{CO}_2]$ (c_i) to c_a : $A/g = c_a (1 - c_i/c_a)/1.6$, where 1.6 is the ratio of gaseous diffusivities of CO_2 and water vapour in air. The c_i/c_a was determined from the stable C isotope compositions of leaves.

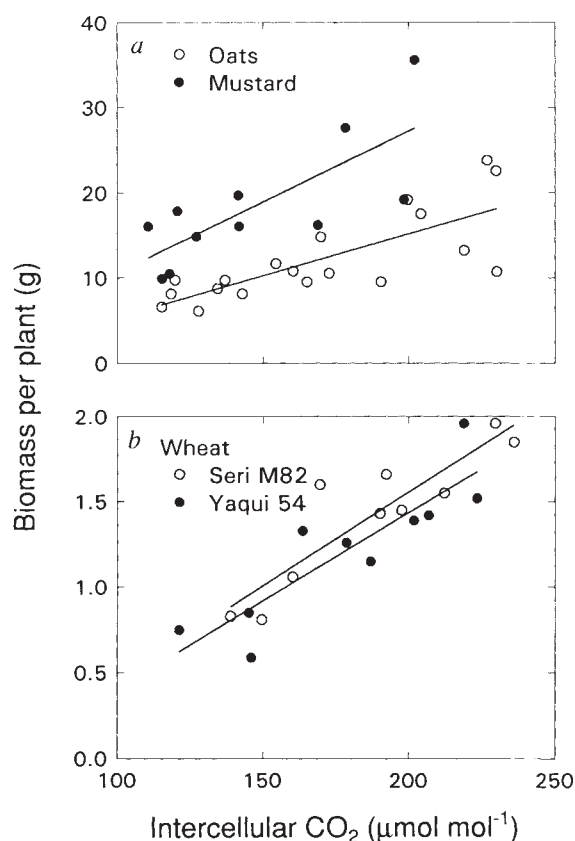


FIG. 4 Relationships of above-ground biomass per C3 plant at maturity to leaf intercellular $[\text{CO}_2]$ (c_i). Leaf c_i was determined from the stable C isotope composition of plants grown from current atmospheric $[\text{CO}_2]$ to *a*, near $150 \mu\text{mol mol}^{-1}$ (oats, mustard; 20 plants m^{-2}) or *b*, $200 \mu\text{mol mol}^{-1}$ CO_2 (wheat; 350 plants m^{-2}). Note that the scale of the ordinate differs in *a* and *b*. Lines are linear regressions of above-ground biomass per plant on leaf c_i : biomass = $-4.42 + 0.10c_i$ for oats ($N = 19$) and biomass = $-6.00 + 0.17c_i$ for mustard ($N = 11$), $r^2 = 0.58$, $P < 0.01$ and biomass = $-0.63 + 0.01c_i$, $r^2 = 0.81$ and biomass = $-0.62 + 0.01c_i$, $r^2 = 0.77$ for Seri M82 and Yaqui 54 cultivars of wheat, respectively, $P < 0.001$, $N = 10$.

understood within an ecosystem context and over relevant periods. Our results nonetheless demonstrate the risks inherent in using present vegetation-climate relationships to reconstruct past climates from pollen or fossil records without incorporating potential direct effects of $[\text{CO}_2]$ ²⁶.

Received 27 April; accepted 10 November 1992.

- Delmas, R. J., Ascencio, J.-M. & Legrand, M. *Nature* **284**, 155-157 (1980).
- Stuiver, M., Burk, R. L. & Quay, P. D. *J. geophys. Res.* **89**, 11713-11748 (1984).
- Neftel, A., Moore, E., Oeschger, H. & Stauffer, B. *Nature* **315**, 45-47 (1985).
- Keeling, C. D. *et al.* in *Aspects of Climate Variability in the Pacific and the Western Americas* (ed. Peterson, D. H.) 165-236 (American Geophysical Union, Washington DC, 1989).
- Henderson, S. A., von Caemmerer, S. & Farquhar, G. D. *Aust. J. Plant Physiol.* **19**, 363-386 (1992).
- Marino, B. D. & McElroy, M. B. *Nature* **349**, 127-131 (1991).
- Masle, J., Farquhar, G. D. & Gifford, R. M. *Aust. J. Plant Physiol.* **17**, 465-487 (1990).
- Wong, S.C., Cowan, I. R. & Farquhar, G. D. *Nature* **282**, 424-426 (1979).
- Polley, H. W., Johnson, H. B. & Mayeux, H. S. *Inter. J. Plant Sci.* **153**, 453-461 (1992).
- Woodward, F. I. *Climate and Plant Distribution* (Cambridge University Press, UK, 1987).
- Stephenson, N. L. *Am. Nat.* **135**, 649-670 (1990).
- Farquhar, G. D. & Richards, R. A. *Aust. J. Plant Physiol.* **11**, 539-552 (1984).
- Morison, J. I. L. *Plant Cell Environ.* **8**, 467-474 (1985).
- Ehleringer, J. R. & Cooper, T. A. *Oecologia (Berlin)* **76**, 562-566 (1988).
- Baker, J. T., Allen, L. H. Jr & Boote, K. J. *J. Agric. Sci., Camb.* **115**, 313-320 (1990).
- Allen, L. H. Jr, Bisbal, E. C., Boote, K. J. & Jones, P. H. *Agron. J.* **83**, 875-883 (1991).
- Bazzaz, F. A. & Fajer, E. D. *Sci. Am.* **266**, 68-74 (1992).
- Shaver, G. R. *et al.* *BioScience* **42**, 433-441 (1992).
- Grulke, N. W., Riechers, G. H., Oechel, W. C., Hjeltn, U. & Jaeger, C. *Oecologia (Berlin)* **83**, 485-494 (1990).
- Norby, R. J., Gunderson, C. A., Wullschlegel, S. D., O'Neill, E. G. & McCracken, M. K. *Nature* **357**, 322-324 (1992).

21. Ehleringer, J. & Björkman, O. *Plant Physiol.* **59**, 86–90 (1977).
22. La Marche, V. C. Jr, Graybill, D. A., Fritts, H. C. & Rose, M. R. *Science* **225**, 1019–1021 (1984).
23. Tans, P. P., Fung, I. Y. & Takahashi, T. *Science* **247**, 1431–1438 (1990).
24. Kauppi, P. E., Mielikäinen, K. & Kuusela, K. *Science* **256**, 70–74 (1992).
25. Kohlmaier, G. H. *et al. Tellus* **41B**, 487–510 (1989).
26. Johnson, H. B., Polley, H. W. & Mayeux, H. S. *Vegetatio* (in the press).
27. Mayeux, H. A., Johnson, H. B., Polley, H. W., Dumesnil, M. J. & Spänel, G. A. *Funct. Ecol.* (in the press).
28. Farquhar, G. D., O'Leary, M. H. & Berry, J. A. *Aust. J. Plant Physiol.* **9**, 121–137 (1982).
29. Craig, H. J. *Geol.* **62**, 115–149 (1954).
30. Roeske, C. A. & O'Leary, M. H. *Biochemistry* **23**, 6275–6284 (1984).

ACKNOWLEDGEMENTS. We thank A. Gibson for preparing samples and figures, B. B. McInteer for measuring C isotopes in plants, and W. A. Dugas, S. W. Leavitt and L. Sternberg for critically reading the manuscript.

Primitive dinosaur skeleton from Argentina and the early evolution of Dinosauria

Paul C. Sereno*, Catherine A. Forster*,
Raymond R. Rogers† & Alfredo M. Monetta‡

* Department of Organismal Biology and Anatomy, University of Chicago, 1027 East 57th Street, Chicago, Illinois 60637, USA

† Department of Geophysical Sciences, University of Chicago, 5734 South Ellis Avenue, Chicago, Illinois 60637, USA

‡ Museo de Ciencias Naturales, Universidad Nacional de San Juan, San Juan 5400, Argentina

WE report here the discovery of a primitive dinosaur skeleton from Upper Triassic strata in northwestern Argentina. The 1-m-long skeleton represents a new taxon, *Eoraptor lunensis* gen. et sp. nov., which is close to the predicted structure and size of the common dinosaurian ancestor^{1–5}. The skull, which has a unique heterodont dentition, does not exhibit any of the specializations of the major dinosaurian clades (Ornithischia, Sauropodomorpha, Theropoda). The forelimbs are less than half the length of the hind limbs, which suggests an obligatory bipedal posture. Although close in overall form to the common dinosaurian ancestor, the functionally tridactyl, grasping/raking hand and other features show that *Eoraptor* is allied phylogenetically with theropods. The discovery of *Eoraptor* supports the hypothesis that dinosaurs diverged rapidly at small body size from a common ancestor, with the

principal herbivorous and carnivorous lineages present by the middle Carnian.

Skeletal remains of the earliest (middle Carnian) dinosaurs are rare and have been unearthed mostly from the fossiliferous Upper Triassic Ischigualasto Formation in northwestern Argentina. These dinosaurs include the ornithischian *Pisanosaurus*⁶ and the theropod *Herrerasaurus*^{1,7–9}. The new skeleton, a contemporary of *Herrerasaurus* in the lower third of the formation (R.R.R., C. C. Swisher, P.C.S., C.A.F. and A.M.M., manuscript in preparation), represents a third species. Despite its small size, closure of sutures in the vertebral column and partial fusion of the scapulocoracoid suggest that the skeleton had reached adult size.

Order Saurischia

Suborder Theropoda

Eoraptor lunensis gen. et sp. nov.

Etymology. *Eos*, dawn (Greek) in reference to its primitive structure and early temporal occurrence; *raptor*, plunderer (Latin) in reference to its carnivorous habits and grasping hand; *luna*, moon (Latin) and *-ensis*, place (Latin), in reference to the type locality.

Holotype. PVSJ 512 (Museo de Ciencias Naturales, Universidad Nacional de San Juan, San Juan, Argentina), an articulated skeleton lacking only the distal caudal vertebrae (Figs 1 and 2). The skeleton was preserved in a muddy siltstone in close association with many of the common vertebrates in the Ischigualasto fauna (*Herrerasaurus*, *Saurosuchus*, *Aetosauroides*, *Ischigualastia*, *Exaeretodon*, and an undescribed small carnivorous cynodont).

Horizon and locality. Ischigualasto Formation (Upper Triassic: middle Carnian); Valley of the Moon (Ischigualasto Provincial Park, Ischigualasto–Villa Unión Basin, northwestern Argentina) (R.R.R. *et al.*, manuscript in preparation).

Diagnosis. Small 1-m-long theropod with external naris slightly enlarged, premaxilla with slender posterolateral process, and leaf-shaped premaxillary and anterior maxillary crowns.

The skull follows a primitive saurischian design, with transversely narrow proportions, a relatively large antorbital opening, a small subnarial foramen beneath the external naris, and a forked posterior process on the jugal (Fig. 1a, b). Derived cranial characters that distinguish ornithischians, sauropodomorphs, and theropods, however, are absent. For example, there is no

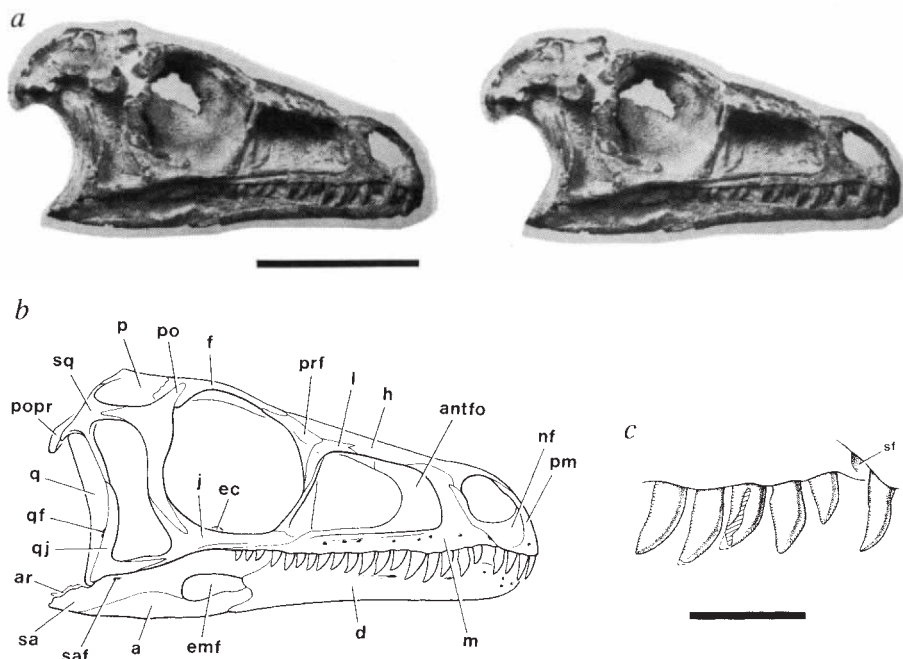


FIG. 1 PVSJ 512, *Eoraptor lunensis*, new species. a, Stereopair and b, reconstruction of the skull in right lateral view. c, Right posterior premaxillary and anterior maxillary teeth in lateral view. Abbreviations: a, angular; antfo, antorbital fossa; ar, articular; d, dentary; emf, external mandibular fenestra; ec, ectopterygoid; f, frontal; j, jugal; l, lacrimal; m, maxilla; n, nasal; nf, narial fossa; p, parietal; pm, premaxilla; po, postorbital; popr, paroccipital process; prf, prefrontal; q, quadrate; qf, quadrate foramen; qj, quadratojugal; sa, surangular; saf, surangular foramen; sf, subnarial foramen; sq, squamosal. Scale bar, a, 5 cm; c, 1 cm.

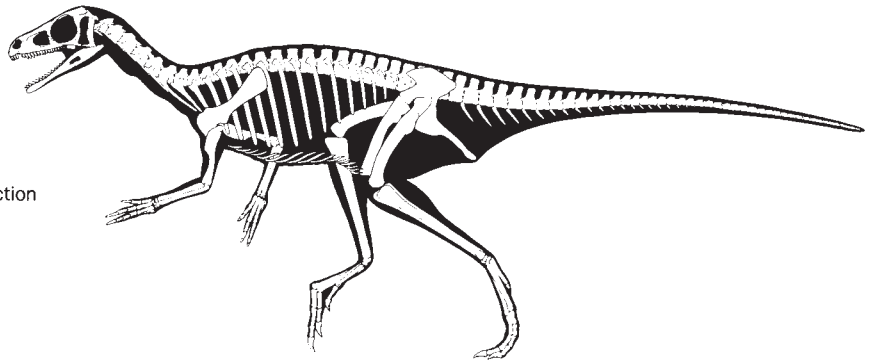


FIG. 2 *Eoraptor lunensis*, new species. Skeletal reconstruction based on PVSJ 512 (length 1 m).

development of an intra-mandibular joint, as occurs in *Herrerasaurus* and other theropods⁹, and the external naris and the narial fossa are not greatly enlarged as in sauropodomorphs¹¹. The dentition is heterodont (Fig. 1b, c). The posterior half of the upper tooth row has serrated, recurved crowns typical of theropods. The crowns in the anterior half of the tooth row, in contrast, are leaf-shaped with a basal constriction that most closely resembles the crown shape in basal sauropodomorphs^{11,12}.

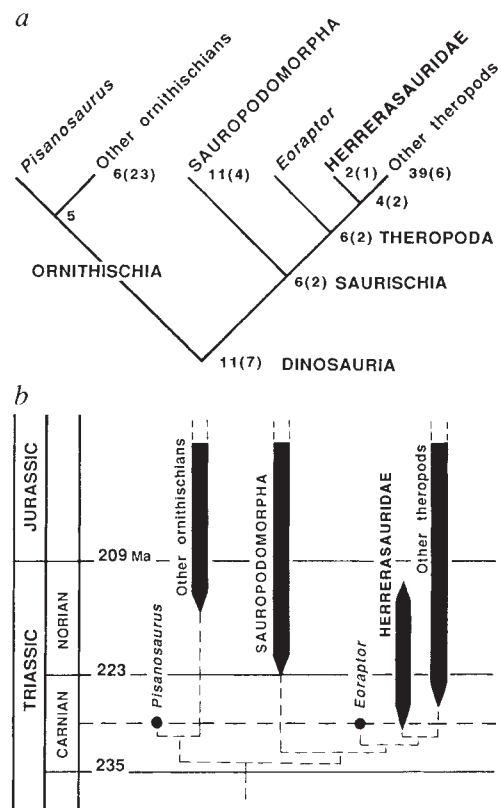
The skeleton exhibits several dinosaurian hallmarks, such as a long deltopectoral crest on the humerus, reduction of the outer digits of the hand (Fig. 2), and an ascending process on the astragalus¹. Similar to the skull, the skeleton lacks many derived characters that distinguish later dinosaurs. There are only three sacral vertebrae that support the pelvis, the lowest number that occurs among dinosaurs, and the hand retains a fifth digit, albeit reduced to a short metacarpal. There are no pneumatic cavities in the presacral vertebrae, as often occurs among saurischians,

but all of the long bones of the skeleton are strongly hollowed as in theropods.

A detailed cladistic analysis of the early branching history of Dinosauria supports a basal split into ornithischians and saurischians, the latter including herrerasaurids (Fig. 3a). Although *Eoraptor* is structurally close to the common dinosaurian ancestor, several advanced features place it among the saurischians as the most primitive known theropod (Fig. 3). Key saurischian synapomorphies (derived characters) present in *Eoraptor* include a subnarial foramen in the skull (Fig. 1c, sf) and modifications that increase the flexibility of the neck (epiphyseal processes and slender, interconnecting ribs). Features that link *Eoraptor* with *Herrerasaurus* and other theropods include a predatory manus with trenchant unguals and enhanced grasping and raking capabilities (elongate distal phalanges, metacarpal pits for phalangeal hyperextension)^{1,2}.

The discovery of *Eoraptor* has opened a window into the early evolution of theropods. The contemporaneous occurrence of a

FIG. 3 a, Hypothesis of early dinosaurian phylogeny based on numerical cladistic analysis of 132 anatomical characters in 12 terminal taxa (consistency index equals 0.85; P.C.S., unpublished data). Numbers at nodes indicate total number of synapomorphies. Synapomorphies that may occupy more than a single node (primarily due to missing data in rare species) are shown in parentheses under delayed transformation¹². Synapomorphies supporting dinosaurian monophyly and internal nodes include: **Dinosauria**, postfrontal absent^{1,2,9}, frontal participation in supratemporal fossa^{1,9}, ectopterygoid overlapping pterygoid⁹, quadrate head exposed laterally^{1,9}, post-temporal opening reduced to a foramen^{1,9}, dorsal vertebra added to sacrum^{1,5,6}, deltopectoral crest 35% or more of humeral length^{1,3-5,13}, manual digit IV narrower than digits II and III and lacking an ungual^{1-3,5,13,14}, acetabular antitrochanter partially divided by notch¹, brevis fossa^{1,14}, femoral head subrectangular^{1-3,5,6,8,15} with angular greater trochanter, femoral medial tuberosity rudimentary, cnemial crest on tibia^{1,2,14}, ascending process on astragalus^{1-6,14,15}, astragalus with laterally facing calcaneal facet, calcaneal medial process rudimentary¹, distal tarsal 4 heel shallow¹, metatarsal 4 shaft sigmoid. **Ornithischia**, subtriangular maxillary/dentary crowns^{2,16}, largest maxillary/dentary tooth in centre of tooth row¹⁶, dentary coronoid process¹⁶, coronoid region half or more as deep as dentary ramus, external mandibular fenestra reduced; **Saurischia**, subnarial foramen, jugal overlaps lacrimal, jugal posterior process forked, epiphyses on mid- and posterior cervical vertebrae², hyposphene-hypantrium articulation in dorsal vertebrae^{2,3}, mid-cervical ribs long and parallel to cervical column, distal carpal 5 absent, phalanx 1 of manual digit I equal to or longer than metacarpal 1^{2,3}, astragalus ascending process wedge-shaped, metatarsals 2-4 with overlapping proximal ends. **Theropoda**, cervical epiphyses prong-shaped, extreme hollowing of centra and long bones², manual digits IV and V vestigial^{2,13}, metacarpals 1-3 with extensor depressions for phalangeal hyperextension, manus more than 50% length of humerus + radius. **Herrerasauridae + other theropods**, intramandibular joint^{1,2,9}, distal caudal prezygapophyses elongate^{2,8}, scapular blade strap-shaped, manual digits I-III with long penultimate phalanges and trenchant unguals², pubic blade at least six times as long as broad, pubic foot¹, proximal end of fibula 75% or more of proximal width of tibia. b, Corresponding phylogram showing the recorded temporal range for each taxon, with dashed line indicating the approximate age (230 Myr; R.R.R. *et al.*, manuscript in preparation) of the dinosaurs from the Ischigualasto Formation (*Pisanosaurus*, *Eoraptor*, *Herrerasaurus*). Solid bars show the earliest record of ornithischians that



are more advanced than *Pisanosaurus*^{17,18}, prosauropod sauropodomorphs^{19,20}, and theropods that are more advanced than *Eoraptor* and *Herrerasaurus*²¹. Herrerasaurids appear first during the Carnian in South America^{1,7,22} and in North America^{23,24}, where they persist into the Norian.

dinosaur as primitive as *Eoraptor* with the ornithischian *Pisanosaurus* and more advanced theropod *Herrerasaurus* supports the hypothesis that dinosaurs diverged rapidly at small body size from a common ancestor during the early Carnian (Fig. 3b). By the middle Carnian, the principal herbivorous and carnivorous lineages of dinosaurs were established. □

Received 26 November; accepted 16 December 1992.

1. Sereno, P. C. & Novas, F. E. *Science* **258**, 1137–1140 (1992).
2. Gauthier, J. *Mem. Calif. Acad. Sci.* **8**, 1–55 (1986).
3. Bakker, R. T. & Galton, P. M. *Nature* **248**, 168–172 (1974).
4. Brinkman, D. B. & Sues, H.-D. *Palaeontology* **30**, 493–503 (1987).
5. Benton, M. J. in *The Dinosauria* (eds Weishampel, D. B., Dodson, P. & Osmólska, H.) 11–30 (University of California Press, Berkeley, 1990).
6. Bonaparte, J. F. *J. Paleont.* **50**, 808–820 (1976).
7. Reig, O. A. *Ameghiniana* **3**, 2–20 (1963).
8. Novas, F. E. *J. vert. Paleont.* (in the press).
9. Sereno, P. C. & Novas, F. E. *J. vert. Paleont.* (in the press).
10. Huene, F. von *Geol. Paläontol. Abh.* **15**, 129–179 (1926).
11. Cooper, M. R. *Occas. Pap. natn. Mus. Rhodesia (B) Nat. Sci.* **6**, 689–840 (1981).
12. Swofford, D. L. *PAUP 3.0* (Illinois Natural History Survey, Champaign, 1989).
13. Sereno, P. C. *J. vert. Paleont.* (in the press).
14. Gauthier, J. & Padian, K. in *The Beginnings of Birds* (eds Hecht, M. K., Ostrom, J. H., Viohl, G. & Wellnhofer, P.) 185–197 (Freunde des Jura-Museums Eichstätt, Eichstätt, 1985).
15. Novas, F. E. *Palaeontology* **35**, 51–62 (1992).
16. Sereno, P. C. *Natn. Geogr. Research* **2**, 234–256 (1986).
17. Chatterjee, S. *Naturwissenschaften* **71**, 630–631 (1984).
18. Hunt, A. P. in *Dawn of the Age of Dinosaurs in the American Southwest* (eds Lucas, S. G. & Hunt, A. P.) 355–358 (New Mexico Museum of Natural History, Albuquerque, 1989).
19. Galton, P. M. *Stuttgarter Beitr. Naturk.* **106**, 1–25 (1984).
20. Galton, P. M. in *The Dinosauria* (eds Weishampel, D. B., Dodson, P. & Osmólska, H.) 320–344 (University of California Press, Berkeley, 1990).
21. Padian, K. in *The Beginning of the Age of Dinosaurs* (ed. Padian, K.) 45–60 (Cambridge Univ. Press, New York, 1986).
22. Colbert, E. H. *Am. Mus. Novitates* **2405**, 1–39 (1970).
23. Murry, P. A. & Long, R. A. in *Dawn of the Age of Dinosaurs in the American Southwest* (eds Lucas, S. G. & Hunt, A. P.) 29–64 (New Mexico Museum of Natural History, Albuquerque, 1989).
24. Parrish, J. M. in *Dawn of the Age of Dinosaurs in the American Southwest* (eds Lucas, S. G. & Hunt, A. P.) 360–374 (New Mexico Museum of Natural History, Albuquerque, 1989).

ACKNOWLEDGEMENTS. We thank C. Abraczinskas for the finished illustrations, W. Simpson and R. Masek for preparation of the skeleton, and M. Carrano, R. Blob, H.-D. Sues, J. Wilson and C. Yu for review of the manuscript. This research was supported by grants from The David and Lucile Packard Foundation, the Petroleum Research Fund of the American Chemical Society, National Geographic Society and NSF (to P.C.S.).

Molecular identification of microorganisms associated with parthenogenesis

R. Stouthamer^{*†}, J. A. J. Breeuwer[†], R. F. Luck[‡] & J. H. Werren[†]

^{*} Department of Entomology, Agricultural University, PO Box 8031, 6700 EH Wageningen, The Netherlands

[†] Department of Biology, University of Rochester, Rochester, New York 14627, USA

[‡] Department of Entomology, University of California, Riverside, California 92521, USA

CYTOPLASMICALLY inherited microorganisms are widespread in insects and have been implicated as causes of female parthenogenesis (females developing from unfertilized eggs) and cytoplasmic incompatibility^{1–15}. Normal sexual reproduction can be restored by treatment with antibiotics^{1–4}. Sequence analysis of the DNA encoding 16S ribosomal RNA has shown that cytoplasmic incompatibility bacteria from diverse insect taxa are closely related (they share >95% sequence similarity) and belong to the alpha subdivision of Proteobacteria^{5–7}. Here we show that parthenogenesis-associated bacteria from parasitoid Hymenoptera also fall into this bacterial group, having up to 99% sequence similarity to some incompatibility microorganisms. Both incompatibility and parthenogenesis microorganisms alter host chromosome behaviour during early mitotic divisions of the egg^{13–17}. Incompatibility bacteria act by interfering with paternal chromosome incorporation in fertilized eggs, whereas parthenogenesis bacteria prevent segregation of chromosomes in unfertilized eggs. These traits are

adaptive for the microorganisms. On the basis of their sequence similarities, we conclude that parthenogenesis bacteria and cytoplasmic incompatibility bacteria form a monophyletic group of microorganisms that 'specialize' in manipulating chromosome behaviour and reproduction of insects.

Identification of insect bacterial symbionts has long been hampered by the inability to culture these fastidious prokaryotes. But the development of the polymerase chain reaction (PCR) and use of DNA sequence encoding 16S rRNA (16S sequences) in microbial phylogeny¹⁸ has made it possible to determine their phylogenetic position. We amplified and sequenced the bacterial 16S ribosomal DNA from six parthenogenetic strains of *Trichogramma* wasps (three different species) and one parthenogenetic strain of *Muscidifurax uniraptor*. *Trichogramma* wasps are minute (around 0.5 mm) parasites of insect eggs, primarily those of Lepidoptera. *Muscidifurax uniraptor* is a Pteromalid pupal parasitoid of houseflies. Each of these strains harbours microorganisms associated with parthenogenesis. To check for PCR amplification of bacteria not associated with parthenogenesis, sexual strains of *Trichogramma* were used as a control. Some forms of parthenogenesis in *Trichogramma* are genetically based and are not associated with microorganisms^{1,3}. As a second control, three strains of *Trichogramma* with genetic parthenogenesis were also examined for cytoplasmic

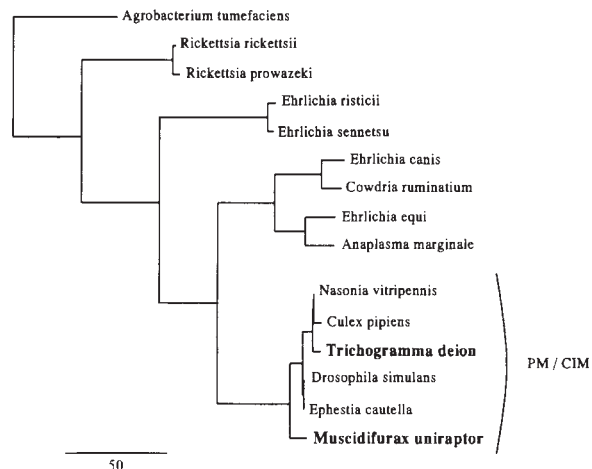


FIG. 1 Most parsimonious phylogenetic tree of parthenogenesis microorganisms (PM) of *Trichogramma deion* (Bautista strain) and *Muscidifurax uniraptor*, several cytoplasmic incompatibility microorganisms (CIM)^{5,6} and other representatives of the alpha subdivision of the Proteobacteria^{21–23}. The 16S rDNA sequences of PM (bold) and CIM are identified by the host species from which they were isolated. *Escherichia coli* (gamma subdivision) was used as an outgroup. Sequences were manually aligned using regions of the 16S gene that are conserved in eubacteria²¹. The aligned sequences of parthenogenesis microorganisms were 1,512 bases in length, including gaps. Gaps were treated as a 'fifth' base. The aligned sequence dataset was analysed with PAUP 3.0 (ref. 20) using the Branch and Bound algorithm to find the shortest tree(s). Two parsimonious trees (length 1,195 bases) were generated. They only differed in the positioning of the *Drosophila simulans* and *Ephestia cautella* microorganisms within the PM/CIM group. This is a result of the fact that only a partial 16S sequence is available for these two cytoplasmic incompatibility microorganisms. Sequences of 16S ribosomal genes of the *M. uniraptor* and *Trichogramma* symbionts were determined using the following procedure. Total DNA was extracted from white pupae (100 of *Trichogramma* spp. and 5 of *Muscidifurax*) after these were surface-sterilized with 70% ethanol, thoroughly washed with sterile water and homogenized in a Mini-Bead beater (Biospec). After precipitation with ethanol, bacterial 16S rDNA was PCR-amplified using conserved 16S rDNA primers (fD2 and rP2; ref. 21). PCR product was then directly cloned into T-tailed M13mp18 vector and sequenced by the Sanger method using Sequenase v. 2.0 kit (US Biochemical; for sequence primers, see refs 5, 27). The nucleotide sequence data of parthenogenesis bacteria will appear in Genbank under the following accession numbers: *M. uniraptor*, L02882; *T. cordubensis*, L02883; *T. deion* TX, L02884; *T. pretiosum*, L02885, *T. deion* Mo, L02886; *T. deion* Ba, L02887; *T. deion* SD, L02888.

TABLE 1 Diagnostic nucleotide positions of subgroups of the parthenogenesis and incompatibility bacteria

Position	Type 1		Type 2
	CIM (4 species)	PM (3 species, 6 strains)	CIM/PM (4 species)
649	A	G	A
650	T	C	T
693	A	A	G
760	A	A	G
844	—	—	T
1,037	A	A	G
1,047	A	A	G
1,127	G	G	A
1,129	T	C	C
1,143	A	A	G
1,144A	T	T	—
1,210	T	T	C
1,245	T	T	G
1,262	G	G	A
1,268	A	A	G
1,285	—	—	T
1,292	A	A	C
1,327	A	A	G
1,364	T	T	C
1,442	G	G	A
1,457	A	A	G

Nucleotide positions of parthenogenesis and incompatibility bacteria are shown that distinguish the two sequence subgroups. Although additional variable positions (47) are present, the positions shown are fixed for all members of each subgroup (so they are diagnostic); position number is based on *E. coli* numbering²⁶. Diagnostic positions for the *Trichogramma* parthenogenesis microorganisms are boxed. CIM, cytoplasmic incompatibility microorganisms; PM, parthenogenesis microorganisms.

microorganisms. For *M. uniraptor*, a closely related sexual species (*M. raptor*) was used as a PCR control. In addition, standard controls for PCR contamination were set up. PCR products were obtained from all microbe-associated parthenogenetic *T. deion* (four different North American collections originating from South Dakota, Texas, and two different localities in California: Bautista Canyon and Mountain Center), *T. pretiosum* (from Mexico), *T. cordubensis* (from Spain) and *M. uniraptor*. No DNA was amplified from the controls.

The results from PCR are consistent with those from cytogenetic and antibiotic studies. Bacteria are easily visualized in the eggs of parthenogenetic *Trichogramma* strains⁴, and antibiotic treatment results in reversion to sexual reproduction and elimination of the microorganisms^{1,4}. Bacteria are not present

in the eggs of bisexual strains or those with genetically based parthenogenesis. All microbe-associated parthenogenetic strains examined have a cytogenetic mechanism of parthenogenesis (gamete duplication) that differs from those with genetic parthenogenesis (ref. 19; R. S. and D. J. Kazmer, unpublished data).

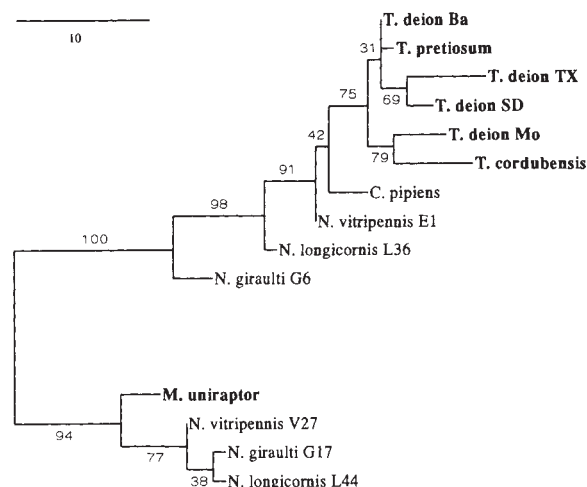
Phylogenetic analysis of the 16S rDNA sequences from the parthenogenesis microorganisms was undertaken using the parsimony algorithm Branch and Bound of PAUP 3.0 (ref. 20). All parthenogenesis bacteria fall in the alpha subdivision of Proteobacteria, and form a group with over 95% sequence similarity (Figs 1 and 2). Most interestingly, all parthenogenesis microorganisms are closely related to cytoplasmic incompatibility bacteria, which are found in diverse insect taxa including beetles (Coleoptera), butterflies (Lepidoptera), flies (Diptera) and wasps (Hymenoptera)⁵⁻⁷. *Wolbachia pipiens*, the cytoplasmic incompatibility microbe of *Culex pipiens*, is type species of the genus *Wolbachia* (see ref. 6). Based on their high sequence similarity, it is reasonable also to place parthenogenesis bacteria within this genus.

Figure 1 shows the phylogenetic position of the parthenogenesis/incompatibility group relative to several representatives of the alpha subdivision of proteobacteria. The parthenogenesis/incompatibility (*Wolbachia*) bacteria form a monophyletic group relative to other alpha Proteobacteria. The closest sequenced relatives to this group are *Anaplasma marginale*²¹, *Cowdria ruminantium*²², and several species of *Ehrlichia*²³. Each of these are arthropod vectored microorganisms that cause mammalian diseases.

A more detailed phylogenetic analysis of the parthenogenesis/incompatibility group is shown in Fig. 2. Although partial 16S sequences (the first 700 bases) are known for incompatibility bacteria from several insect species^{6,7}, nearly complete sequences have only been published for incompatibility bacteria of the mosquito *Culex pipiens*⁶ and wasps of the genus *Nasonia*⁵. This is unfortunate, because much of the phylogenetically informative sequence information is in the second half (3' end) of the 16S rDNA gene⁵. Therefore we did the phylogenetic analysis for those incompatibility microbe sequences that are nearly complete, including *Culex*⁶ and *Nasonia* incompatibility bacteria⁵.

The parthenogenesis/incompatibility group is divided into two subgroups that differ consistently from each other at 15 positions over 1,464 bases of the almost complete 16S ribosomal gene. The two subgroups represent real 16S variants based on analysis of secondary structure⁵. The diagnostic base positions separating the two subgroups are shown in Table 1. Although incompatibility microorganisms from different insects occur in both subgroups, all parthenogenesis microorganisms from

FIG. 2 A most parsimonious phylogenetic tree of parthenogenesis and cytoplasmic incompatibility microorganisms^{5,6} based on 16S rDNA sequences. The 16S sequences of microorganisms are identified by the host from which they are isolated (PM are in bold). *Anaplasma marginale* was used as the outgroup. Only taxa with complete 16S rDNA sequences were included in the analysis. Gaps were treated as a 'fifth' base. Constant and uninformative sites were eliminated, which gave a total of 68 informative (variable) sites. The reduced dataset was analysed using PAUP 3.0 and the Branch and Bound algorithm²⁰. Numbers indicate the level of support for individual nodes on the tree based on 100 bootstrapped (Branch and Bound) runs. Twelve most parsimonious trees (length 97) were generated with minor differences at three nodes, which is consistent with the low levels of support for those particular nodes. Sequences from each subgroup have been identified from each *Nasonia* species. These represent either two 16S rDNA genes in *Nasonia* *Wolbachia* or infection by two different bacterial strains⁵.



Trichogramma wasps fall into subgroup I. All 16S rDNA sequences of parthenogenesis bacteria from *Trichogramma* can be distinguished from the incompatibility bacteria by two diagnostic base positions, as shown in Table 1. This suggests a monophyletic origin for *Trichogramma* parthenogenesis bacteria. In contrast, the parthenogenesis bacterium of *Muscidifurax* falls into subgroup II, suggesting an independent derivation of bacterial parthenogenesis in this wasp. Alternatively, microbial parthenogenesis may have a single origin, with subsequent sequence divergence. Resolving such issues requires more sequence information on hosts and bacterial symbionts.

Both parthenogenesis and incompatibility increase the transmission of these cytoplasmically (maternally) inherited bacteria. Parthenogenesis increases frequency of the bacteria by biasing sex ratio towards the transmitting (female) sex. Incompatibility increases the frequency of associated bacteria by indirectly decreasing the frequency of cytoplasmic lineages that do not carry the same bacterial strain^{12,24}. Thus these phenotypes are selectively advantageous for the microorganisms.

The parthenogenesis/incompatibility *Wolbachia* are cytoplasmically inherited symbionts of insects that have evolved mechanisms of altering mitosis in their insect hosts. Cytoplasmic incompatibility bacteria have been shown to disrupt the first mitotic divisions in fertilized eggs of incompatible crosses in *Nasonia*^{13,14}, *Drosophila simulans*¹⁵ and *Culex pipiens*¹⁶. In *Nasonia* it has further been shown that condensation of chromosomes derived from the sperm is abnormal and that they are often fragmented in incompatible crosses^{13,14,17}. Mechanisms are unknown, although the process is likely to involve chromosome imprinting^{14,25}. Parthenogenesis bacteria alter the segregation patterns of chromosomes in unfertilized eggs (ref. 19; R.S. and D. J. Kazmer, unpublished data). By preventing segregation of chromosomes in the first mitotic division, diploidy is restored, which leads to the development of a parthenogenetic female. Apparently some special attributes of this microbial group have led to acquisition of the ability to manipulate mitosis in eukaryotic hosts. □

Received 8 June; accepted 21 October 1992.

1. Stouthamer, R., Luck, R. F. & Hamilton, W. D. *Proc. natn. Acad. Sci. U.S.A.* **87**, 2424–2427 (1990).
2. Zchori-Fein, E., Roush, R. T. & Hunter, M. S. *Experientia* **48**, 102–105 (1992).
3. Stouthamer, R., Pinto, J. D., Platner, G. R. & Luck, R. F. *Ann. Entomol. Soc. Amer.* **83**, 475–581 (1990).
4. Stouthamer, R. & Werren, J. H. *J. Invert. Path.* (in the press).
5. Breeuwer, J. A. J. *et al. Insect molec. Biol.* **1**, 25–36 (1992).
6. O'Neill, S. L., Giordano, R., Colbert, A. M. E., Karr, T. L. & Robertson, H. M. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2699–2702 (1992).
7. Rousset, F., Vautrin D. & Solignac, M. *Proc. R. Soc. B* **247**, 163–168 (1992).
8. Yen, J. H. & Barr, A. R. *J. Invert. Path.* **22**, 242–250 (1973).
9. Kellen, W. R., Hoffmann, D. F. & Kwock, R. A. *J. Invert. Path.* **37**, 273–283 (1981).
10. Wade, M. J. & Stevens, L. *Science* **278**, 527–528 (1985).
11. Hoffmann, A. A., Turelli, M. & Simmons, G. M. *Evolution* **40**, 692–701 (1986).
12. Richardson, P. M., Holmes, W. P. & Saul, G. B. *J. Invert. Path.* **50**, 176–183 (1987).
13. Ryan, S. L. & Saul, G. B. *Molec. gen. Genet.* **103**, 29–36 (1968).
14. Breeuwer, J. A. J. & Werren, J. H. *Nature* **346**, 558–560 (1990).
15. O'Neill, S. L. & Karr, T. L. *Nature* **348**, 178–180 (1990).

16. Jost, E. *Wilhelm Roux Arch. entwicklungsmech. Org.* **166**, 173–188 (1970).
17. Ryan, S. L., Saul, G. B. & Conner, G. W. *J. Hered.* **76**, 21–26 (1985).
18. Woese, C. R. *Microbiol. Rev.* **51**, 221–271 (1987).
19. Legner, E. F. *Can. Entomol.* **117**, 383–389 (1985).
20. Swofford, D. L. PAUP V3.0 Illinois Natural History Survey, Champaign, IL (1990).
21. Weisburg, W. G., Barns, S. M., Pelletier, D. A. & Lane, D. J. *J. Bact.* **173**, 697–703 (1991).
22. van Vliet, A. H. M., Jongejans, F. & van der Zeijst, B. A. M. *Int. J. Syst. Bact.* **42**, 494–502 (1992).
23. Anderson, B. E., Dawson, J. E., Jones, D. C. & Wilson, K. H. *J. clin. Microbiol.* **29**, 2838–2842 (1991).
24. Caspari, E. & Watson, G. S. *Evolution* **13**, 568–570 (1959).
25. Jablonka, E. & Lamb, M. J. *J. theor. Biol.* **139**, 69–83 (1989).
26. Brosius, J., Palmer, J. L., Kennedy, J. P. & Noller, H. F. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4801–4805 (1978).
27. Lane, D. J. in *Nucleic Acid Techniques in Bacterial Systematics* (eds Stackebrandt, E. & Goodfellow, M.) 115–174 (Wiley, Chichester, 1991).

ACKNOWLEDGEMENTS. We thank J. D. Pinto and G. R. Platner for *Trichogramma* cultures, E. F. Legner for *Muscidifurax* cultures, and W. Burke and T. Eickbush for advice. This work was supported by USDA.

FGF-4 and BMP-2 have opposite effects on limb growth

Lee Niswander & Gail R. Martin

Department of Anatomy and Program in Developmental Biology,
School of Medicine, University of California, San Francisco,
California 94143, USA

LIMB development is dependent on epithelial–mesenchymal interactions. The apical ectodermal ridge (AER), a specialized epithelium at the limb tip, stimulates proliferation of underlying mesenchyme, causing directed limb outgrowth¹ (for review see ref. 2). Several genes are expressed in the mouse AER^{3–10}, including *Fgf-4* (fibroblast growth factor-4)^{11,12} and *Bmp-2* (bone morphogenetic protein-2)¹³, both of which encode secreted signalling molecules. Using a culture system developed to explore the function of molecules produced by the AER, we have shown that FGF-4 protein stimulates proliferation of mesenchyme in the early mouse limb-bud. This suggests that FGF-4 serves that major function of the AER. In contrast, BMP-2 inhibits limb growth, suggesting that as a result the AER may serve a hitherto unrecognized inhibitory function. Furthermore, the extent of limb outgrowth can be modulated by mixing the two signalling molecules, suggesting that limb growth is regulated by a combination of stimulatory and inhibitory signals from the AER.

In the culture system used (see legend to Fig. 1), trunk fragments are isolated containing either forelimb or hindlimb buds, and the AER is removed from one limb (–AER limb) to expose the underlying mesenchyme, but kept intact on the other

(+AER limb). These are cultured in serum-free defined medium with or without growth factor supplementation for up to 48 h. This system allows direct comparison of contralateral limbs from a given embryo, controlling for variability among embryos and culture conditions. Under these conditions AER regeneration does not occur, and in contrast to what has been observed in the chick¹⁴, there is no obvious effect of AER removal on limb outgrowth (Figs 1a, b and 2a). AER removal does, however, cause a decrease (~45%) in the mitotic index in distal mesenchyme (Fig. 3). Cell morphology is not significantly affected, and pyknotic cells and cell debris are observed only in the outermost layer of exposed mesenchyme. Cells beneath this layer appear healthy, showing little evidence of cell death (Fig. 2d, e).

Fgf-4 (previously known as *hst-1* or *kFGF*), which encodes a secreted protein FGF-4 (ref. 15), is expressed in the AER during the period when that tissue mediates the proliferative response of the underlying mesenchyme^{11,12}. Addition of recombinant FGF-4 protein to the limb cultures results in a large increase in proximo–distal (P–D) length of the –AER limb as compared to the contralateral +AER limb (Fig. 1c, d), or to limbs cultured without FGF-4 (Fig. 1a, b). Effects are detected within 12 h (data not shown). After 48 h and over a range of FGF-4 concentrations (10, 50, 100 or 500 ng ml^{–1}, about 0.8–40 nM), the average length of –AER forelimbs is 50% greater than that of contralateral +AER forelimbs or forelimbs cultured without FGF-4. Hindlimbs exhibit >100% increase in P–D length compared with +AER hindlimbs or hindlimbs cultured without FGF-4 (Table 1). The greater responsiveness of hindlimbs may reflect greater sensitivity of developmentally younger hindlimb mesenchyme to stimulation by FGF-4, or

alternatively, might be due to differences in mesenchyme specific to hind versus forelimbs.

FGF-4 treatment causes elongation along the P-D axis of -AER limb tissue contained within surrounding epithelium, and extrusion of exposed limb mesenchyme, but has no effect on +AER limbs (Figs 1c, d and 2b, f; Table 1). This is presumably because the intact epithelium acts as a barrier, preventing FGF-4 from reaching the underlying mesenchyme. The observed increase in -AER limb length results from increased cell division in distal limb mesenchyme, as demonstrated by the finding that the frequency of bromodeoxyuridine(BrdU)-labelled mesenchyme cells is 33% ($\pm 4.5\%$) in FGF-4 stimulated -AER limbs, but only 4% ($\pm 0.7\%$) in contralateral +AER limbs (method described in the legend to Fig. 3). Thus exposure to FGF-4 causes a significant increase in the number of cells initiating DNA synthesis. Furthermore, in FGF-4-treated cultures the mitotic index in mesenchyme at the exposed surface and in deeper distal mesenchyme of -AER limbs is 44% higher than in corresponding mesenchyme underlying the AER in contralateral limbs, and 145% higher than in -AER limbs cultured in the absence of FGF-4 (Fig. 3). In contrast, in proximal limb mesenchyme there is no difference in mitotic index between FGF-4-treated -AER and +AER limbs (data not shown). Neither recruitment of cells from adjacent trunk mesenchyme

TABLE 1 Effects of growth factors on P-D length of +AER and -AER limbs

Medium supplement	n	Length of limb in mm (mean $\pm$ s.d.)			
		Forelimbs		Hindlimbs	
		+AER	-AER	+AER	-AER
None	>50	0.70 $\pm$ 0.05	0.70 $\pm$ 0.05	0.40 $\pm$ 0.05	0.35 $\pm$ 0.05
FGF-4†	>50	0.75 $\pm$ 0.15	1.10 $\pm$ 0.15*	0.35 $\pm$ 0.10	0.75 $\pm$ 0.10*
FGF-4, without					
heparan sulphate§	6	0.55 $\pm$ 0.05	0.70 $\pm$ 0.05	0.35 $\pm$ 0.05	0.35 $\pm$ 0.05
BMP-2‡	25	0.65 $\pm$ 0.05	0.50 $\pm$ 0.00†	0.40 $\pm$ 0.10	0.20 $\pm$ 0.10
TGF- β 1	15	0.70 $\pm$ 0.10	0.55 $\pm$ 0.10	0.40 $\pm$ 0.10	0.30 $\pm$ 0.10
BMP-4¶	16	0.65 $\pm$ 0.10	0.65 $\pm$ 0.10	0.40 $\pm$ 0.10	0.40 $\pm$ 0.05
PDGF¶, insulin#					
EGF¶, retinoic acid**	12	0.70 $\pm$ 0.10	0.70 $\pm$ 0.10	0.45 $\pm$ 0.05	0.45 $\pm$ 0.05
FGF-4: BMP-2	12	0.80 $\pm$ 0.05	1.10 $\pm$ 0.10*	0.45 $\pm$ 0.10	0.70 $\pm$ 0.15*
100:10 ng ml ⁻¹					
FGF-4: BMP-2					
50:50 ng ml ⁻¹	16	0.75 $\pm$ 0.05	0.75 $\pm$ 0.05	0.45 $\pm$ 0.05	0.40 $\pm$ 0.05
FGF-4: BMP-2					
10:100 ng ml ⁻¹	10	0.65 $\pm$ 0.05	0.45 $\pm$ 0.05†	0.40 $\pm$ 0.05	0.15 $\pm$ 0.05

Trunk fragments with attached limbs (+AER and -AER) were cultured as described in the legend to Fig. 1 for 48 h in serum-free medium supplemented as stated (resupplemented after 24 h). Measurements of the P-D length, from the distal tip to the point of attachment with the body, were made for each pair of limbs. Lengths are expressed as the mean of the distances measured in millimeters, $\pm$ the standard deviation (s.d.). Growth factors: mature murine FGF-4 protein lacking the secretory signal sequence was produced using the pGEX2T bacterial expression system²⁵. The expression plasmid was constructed by cloning a fragment of *Fgf-4* complementary DNA²⁶ obtained by *EcoRI* and partial *RsrII* digestion into the *EcoRI* site of pGEX2T. Bacterial culture, isopropyl- β -D-thiogalactoside (IPTG) induction (4 h), affinity purification of the fusion protein by absorption to glutathione-agarose beads and thrombin cleavage to release FGF-4 were as described in ref. 25. Most of the free FGF-4 remained bound to beads, allowing removal of residual thrombin by washing. FGF-4 protein was eluted with 1.5 M KCl and dialysed. FGF proteins (FGF-1 provided by L. T. Williams, FGF-2 purchased from Boehringer-Mannheim, FGF-4 produced as described, and FGF-5 provided by J. Hébert) were diluted in 25 mM HEPES and 0.1% BSA (acetylated form; New England Biolabs). Unless otherwise stated, whenever FGFs were added to culture medium, heparan sulphate (10–100 ng ml⁻¹; Sigma) was also added for protein stabilization. Control experiments showed that addition of heparan sulphate, alone or in conjunction with other growth factors, BSA alone, or supernatant prepared appropriately from bacterial cells containing only pGEX2T expression vector, had no apparent effect on limb growth. Recombinant human BMP-2 and BMP-4, generously provided by Genetics Institute (Cambridge, MA), and TGF- β 1 by R. Derynck, were diluted in 20 mM sodium acetate, pH 5.0 and 0.4% BSA. Recombinant human PDGF (Collaborative Biomedical Products) was diluted in 10 mM acetic acid. Murine EGF (Collaborative Biomedical Products) and bovine insulin (Boehringer-Mannheim) were resuspended as specified by the manufacturers. All protein dilutions were prepared in siliconized tubes and used immediately. All-trans-retinoic acid (Sigma) was dissolved in dimethyl sulphoxide (DMSO) (final concentration 0.01%). n, Number of $\pm$ AER forelimb or hindlimb pairs analysed.

* Mesenchymal extrusion (0.25 $\pm$ 0.05 mm).

† Distance in mm along the A-P axis at the distal tip of bud (+AER: 0.45 $\pm$ 0.05; -AER: 0.20 $\pm$ 0.05). Similar results obtained with the following concentrations: ‡ 10, 50, 100 ng ml⁻¹; § 100, 200 ng ml⁻¹; || 50, 100 ng ml⁻¹; ¶ 50, 200 ng ml⁻¹; # 10, 200, 5,000 ng ml⁻¹; ** 1, 10, 100 ng ml⁻¹.

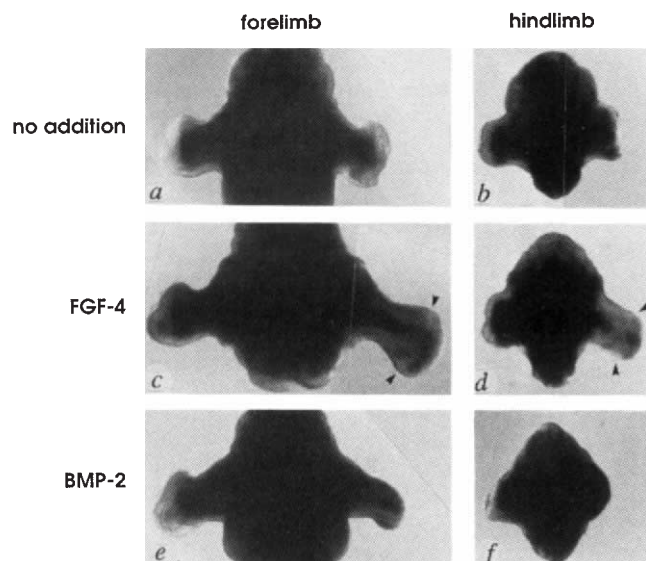


FIG. 1 Effects of FGF-4 and BMP-2 on cultured limbs. Trunk fragments with attached limbs (+AER and -AER) were cultured for 48 h (a, b) without growth factor supplementation; c, d, with 50 ng ml⁻¹ FGF-4; e, f, with 50 ng ml⁻¹ BMP-2. +AER limb is on left and -AER limb on right. Arrowheads (c, d) indicate the distal limit of epithelium covering the -AER limb; extrusion of mesenchyme is visible distally.

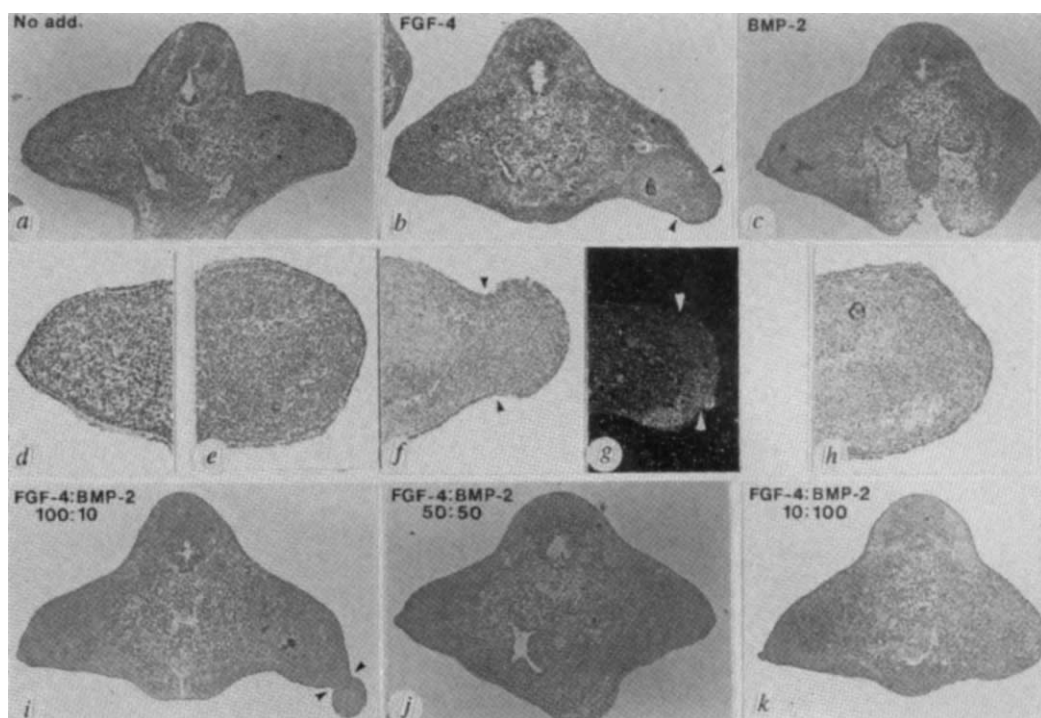
METHODS. Limb staging was as described in ref. 27. Mouse embryos with forelimbs at stage 3–4 (fully formed AER) and hindlimbs at stage 2–3 (apical epithelium thickening) were dissected into cold phosphate-buffered saline containing 10% fetal calf serum, and the heart removed. Limb regions were isolated by four transverse cuts through the trunk, one above and one below each pair of limbs. With fine forceps, the entire AER was teased away from the left limb and discarded, causing no obvious damage to underlying mesenchyme. Trunk fragments, with attached -AER and +AER limbs, were transferred to tubes (Falcon no. 2057; 4–6 trunk fragments per tube) containing 1-ml aliquots of equilibrated serum-free defined culture medium (modified Biggers medium²⁸), growth factors added (see legend to Table 1), and the tubes rotated (30 r.p.m. at $\sim 75^\circ$ angle) in a 5% CO₂ incubator at 37 °C. Trunk tissue appeared similar in medium with or without growth factor supplementation. Over 48 h of culture, development is normal although slightly retarded. Forelimbs develop to stage 6–7 versus stage 8 *in vivo*. Hindlimbs reach stage 5 versus stage 5–6 *in vivo*. Limb tissue appears healthy with little cell death, whereas the trunk region shows signs of deterioration. Limbs cultured for longer periods show a considerable decrease in growth rate and marked tissue degeneration.

nor oedema are responsible for the observed increase in -AER limb length, because similar results are obtained with cultures of detached limbs, and there is no obvious change in mesenchyme cell density following FGF-4 treatment (Fig. 2e, f, and data not shown). Thus, FGF-4 causes limb outgrowth by stimulating proliferation of cells in distal limb mesenchyme. No obvious regional differences in response to ligand were detected in these cultures. This is in contrast to previous findings³¹, in which cultures of dissociated anterior mesenchyme were shown to be more responsive to FGF than were cultures of posterior mesenchyme.

No effects on limb growth are observed when unrelated growth factors (platelet-derived growth factor, epidermal growth factor, insulin) or retinoic acid are added to the culture medium, or in control cultures (Table 1). But other FGF proteins can stimulate limb outgrowth: FGF-1 (acidic FGF), FGF-2 (basic FGF) and FGF-5 have essentially the same effect as FGF-4 (data not shown). Growth stimulation by all FGFs tested is not unexpected because different FGF ligands can interact with and activate the same receptor^{16,17}, and FGF receptors are expressed in limb mesenchyme^{18,19}.

FIG. 2 Mixed signals: effects of combinations of FGF-4 and BMP-2 on hindlimb outgrowth. Sections (6 μm) of trunk fragments with attached hindlimbs cultured for 48 h in: *a, d, e*, medium without growth factor supplementation; or medium containing (*b, f, g*) 50 ng ml⁻¹ FGF-4; *c, h*, 50 ng ml⁻¹ BMP-2; *i, j*, 100 ng ml⁻¹ FGF-4 and 10 ng ml⁻¹ BMP-2; *k*, 10 ng ml⁻¹ FGF-4 and 100 ng ml⁻¹ BMP-2. In panels where both limbs are shown, the +AER limb is on left side, -AER limb on right. *d* and *e-h*, +AER and -AER limbs, respectively, at $\sim\times 5$ higher magnification. *g*, Photographed under dark-field illumination, shows hybridization of a *Hoxd-13* antisense probe. *Hoxd-13* RNA is detected in the posterior (bottom) but not the anterior (top) mesenchyme. Similar effects on limb outgrowth and *Hoxd* gene expression were observed in forelimb cultures. Arrowheads indicate the distal limit of the epithelium covering the -AER limb; extrusion of mesenchyme is visible distally.

METHODS. After culture, embryo fragments, as illustrated in *g*, were



processed for RNA *in situ* hybridization²⁹. Probes for *Hoxd-11* and *Hoxd-13* have been described^{21,30} and were provided by D. Duboule.

In these experiments, all mesenchyme underlying the AER is exposed to exogenous FGF-4. *In vivo*, however, it is likely that only posterior mesenchyme is exposed. *Fgf-4* RNA is localized to the posterior half of the AER¹¹, and the affinity of FGF-4 protein for extracellular matrix components presumably prevents it from diffusing very far from the cells that secrete it. Such spatial restriction could account, at least in part, for the greater growth on the posterior than the anterior side of the limb (reviewed in ref. 32).

In addition to its apparent role in controlling limb mesenchyme proliferation, it is conceivable that FGF-4 also functions in the specification of pattern along the anterior-posterior (A-P) axis by regulating the expression of genes that play a role in this process. To explore this possibility we examined the effect of FGF-4 treatment on the expression of two genes, *Hoxd-11* (*Hox-4.6*) and *Hoxd-13* (*Hox-4.8*), involved in the specification of posterior fate in the limb^{20,21}. In the +AER limb, *Hoxd-11* and *Hoxd-13* RNAs were detected only in the posterior distal mesenchyme, in patterns similar to those observed during normal development *in vivo*²¹ (data not shown). In the -AER limb, posterior restriction of these RNAs was also observed. Even in extruded mesenchyme, they were detected only on the posterior side of the extrusion (Fig. 2g, and data not shown). These results demonstrate that exposure to FGF-4 alone is not sufficient to alter the expression in limb mesenchyme of *Hox* genes involved in A-P patterning, such as *Hoxd-11* and *Hoxd-13*. Utilization of the culture system described here has, however, enabled us to demonstrate that FGF-4 alone is sufficient to regulate the expression of *Eux-1*, a member of a different class of homeobox-containing genes (L.N. and G.R.M., manuscript submitted). *Eux-1* function in limb development is under investigation.

BMP-2, a member of the transforming growth factor- β (TGF- β) superfamily, is encoded by a gene expressed in the AER during the same developmental period as *Fgf-4*¹³. In contrast to the effects of FGF-4, addition of BMP-2 to the culture medium

causes inhibition of -AER limb growth (Figs 1e, f and 2c; Table 1). Effects of this protein are observed within 12 h of culture (data not shown). After 24-48 h of culture in BMP-2 at 10, 50 or 100 ng ml⁻¹, -AER forelimbs are significantly shorter (P-D length) and, strikingly, the A-P width is considerably narrower at the distal tip, when compared with contralateral +AER forelimbs (Table 1). Growth inhibition is particularly evident for -AER hindlimbs (Figs 1f and 2c), which are often almost undetectable. At 1 ng ml⁻¹ BMP-2, growth is slightly inhibited and hindlimbs are affected to a greater extent than forelimbs (data not shown).

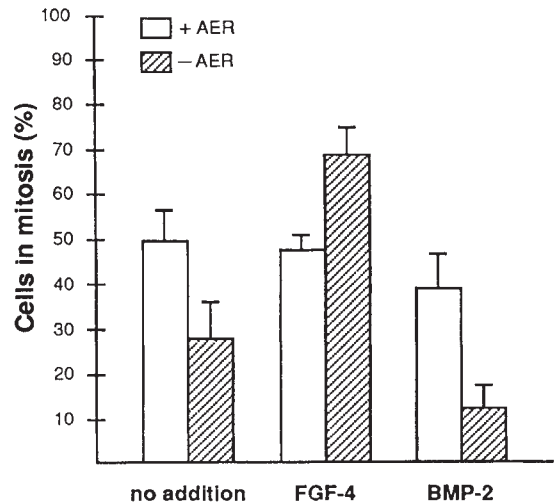
Histological examination of BMP-2-treated -AER limbs revealed pyknotic cells in mesenchyme exposed at the limb surface and to a depth of ~ 4 cell layers. Below this, cells appeared healthy (Fig. 2h, and data not shown). The mitotic index in deeper distal mesenchyme (6-16 cell layers beneath the surface) in -AER limbs is threefold lower than in a comparable region of contralateral +AER limbs, and twofold lower than in -AER limbs cultured in the absence of BMP-2 (Fig. 3), indicating BMP-2 inhibits cell proliferation in distal limb mesenchyme.

In addition, we tested the effect of a closely related molecule, BMP-4, which is also expressed in the AER⁸. Surprisingly, no effect on limb growth was detected with two different preparations of BMP-4 in a total of four experiments at concentrations ranging from 0.1-800 ng ml⁻¹ (Table 1, and data not shown). It is unlikely that these negative results were due to inactivity of BMP-4 protein, because both preparations continued to be active in a different assay²². In contrast, another member of the TGF- β superfamily, TGF- β 1, did inhibit limb growth, but to a lesser extent than BMP-2 (Table 1). These variations in responsiveness to different family members presumably reflect differences in the availability of appropriate receptors, binding proteins, and/or other factors required for ligand activity in the mesenchyme.

We next did a 'mixed signals' experiment, in which limbs

FIG. 3 Effects of FGF-4 and BMP-2 on limb mesenchyme proliferation. The mitotic index (per cent of cells in mitosis) in the distal mesenchyme was determined for +AER and -AER limbs cultured for 28 h in medium alone, or in medium containing FGF-4 or BMP-2 (100 ng ml⁻¹). Nocodazole (5 µg ml⁻¹), which arrests cells in mitosis, was present during the last 6 h of culture.

METHODS. Trunk fragments with attached limbs were fixed in 4% paraformaldehyde, embedded in paraffin, sectioned (6 µm), and stained with haematoxylin and eosin. Using an ocular grid (50 µm × 50 µm), the mitotic index was determined by counting the number of mitoses per total number of nuclei within 50 µm of the AER or exposed surface. For each treatment at least seven pairs of limbs were examined and over 1,500 nuclei counted. Sections analysed were separated by at least 18 µm. The frequency of BrdU-labelled cells was determined according to the following procedure (C. Tickle, personal communication). After 23 h culture in FGF-4 (100 ng ml⁻¹), BrdU (20 µM) was added to the medium and the limbs were cultured for an additional hour. The distal tip (200 µm) of each limb was dissected and placed in 2% trypsin at 4 °C for 30 min to aid removal of the epithelium. Mesenchyme cells were dissociated and smeared on microscope slides, air-dried, fixed in 4% formaldehyde, incubated with anti-BrdU antibody (Boehringer-Mannheim) and secondary fluorescein-tagged antibody. Cells that incorporated BrdU were detected by fluorescence microscopy. The total number of cells and of labelled cells within a randomly selected field of view were counted. Four pairs of limbs were analysed and over 1,500 cells counted for both +AER and -AER limbs.



were cultured in media containing different mixtures of FGF-4 and BMP-2. Supplementation with 100 ng ml⁻¹ FGF-4 and 10 ng ml⁻¹ BMP-2 results in P-D elongation and mesenchyme extrusion indistinguishable from that produced by FGF-4 alone (compare Fig. 2i with b; Table 1). Thus, in the presence of 100 ng ml⁻¹ FGF-4, 10 ng ml⁻¹ BMP-2 has no effect on outgrowth, even though this concentration of BMP-2 alone is sufficient to cause inhibition. The inhibitory effects of BMP-2 are therefore not due to toxicity of the protein preparation. Conversely, a reciprocal combination of the growth factors (10 ng ml⁻¹ FGF-4 and 100 ng ml⁻¹ BMP-2) causes dramatic growth inhibition of -AER limbs, similar to that observed with BMP-2 alone (compare Fig. 2k with c; Table 1). Thus, with 100 ng ml⁻¹ BMP-2, 10 ng ml⁻¹ FGF-4 has no effect on outgrowth, although this concentration of FGF-4 alone is sufficient to stimulate outgrowth. When equal concentrations of FGF-4 and BMP-2 (50 ng ml⁻¹ each) are used, P-D growth is neither stimulated nor inhibited (Fig. 2j; Table 1). In essence, the effect of each growth factor is cancelled by the other, and -AER limbs are similar to those cultured in medium without supplementation (compare Fig. 2j with a; Table 1). This indicates that BMP-2 inhibits the proliferative response to FGF-4 and, conversely, FGF-4 also inhibits the effect of BMP-2, thus providing direct evidence that the opposing effects of FGF-4 and BMP-2 can each be modulated by the presence of the other. Studies aimed at localizing FGF-4 and BMP-2 proteins, as well as other components of their respective signalling pathways, will help to

determine whether such modulation plays a role in limb development *in vivo*.

The importance of the AER in stimulating limb outgrowth has been recognized since the 1940s¹. Our experiments provide, to our knowledge, the first functional analysis of molecules produced in the AER. We show that FGF family members can stimulate limb mesenchyme proliferation, whereas other growth factors (EGF, PDGF, TGF-β superfamily members, insulin) and retinoic acid do not. Because FGF-4 is localized in the AER when that tissue stimulates mesenchyme proliferation, whereas at least four other FGF genes are not^{23,24} (J.-K. Han and G.R.M., unpublished data), it is likely that FGF-4 serves as an endogenous signal for mesenchyme proliferation and thus performs one of the major functions of the AER *in vivo*. We also demonstrate that some TGF-β superfamily members have the opposite effect, and that BMP-2 functions to inhibit mesenchyme proliferation. Our 'mixed signals' experiments further demonstrate that the effects of FGF-4 and BMP-2 can each be modulated by the other. Because these genes are coexpressed in the AER, this raises the possibility that the extent of P-D growth is regulated, at least in part, through an interplay between stimulatory activity of FGF-4 and inhibitory activity of BMP-2. Changes in the local concentrations of these signalling molecules may alter the balance of positive and negative signals for proliferation, leading to refinements in limb shape and limiting its overall P-D outgrowth. □

Received 5 October; accepted 9 November 1992.

1. Saunders, J. W. Jr. *J. exp. Zool.* **108**, 363-403 (1948).
2. Tabin, C. J. *Cell* **66**, 199-217 (1991).
3. Dollé, P. et al. *Nature* **342**, 702-705 (1989).
4. Hill, R. E. et al. *Genes Dev.* **3**, 26-37 (1989).
5. Robert, B., Sassoon, D., Jacob, B., Gehring, W. & Buckingham, M. *EMBO J.* **8**, 91-100 (1989).
6. Gavin, B. J., McMahon, J. A. & McMahon, A. P. *Genes Dev.* **4**, 2319-2332 (1990).
7. Davis, C. A., Holmyard, D. P., Millen, K. J. & Joyner, A. L. *Development* **111**, 287-298 (1991).
8. Jones, C. M., Lyons, K. M. & Hogan, B. L. M. *Development* **111**, 531-542 (1991).
9. Jackson-Grusby, L., Kuo, A. & Leder, P. *Genes Dev.* **6**, 29-37 (1992).
10. Ruberte, E., Friederich, V., Morriss-Kay, G. & Chambon, P. *Development* **115**, 973-987 (1992).
11. Niswander, L. & Martin, G. R. *Development* **114**, 755-768 (1992).
12. Suzuki, H. R. et al. *Dev. Biol.* **150**, 219-222 (1992).
13. Lyons, K. M., Peltton, R. W. & Hogan, B. L. M. *Development* **109**, 833-844 (1990).
14. Summerbell, D. J. *Embryol. exp. Morph.* **32**, 651-660 (1974).
15. Delli-Bovi, P. et al. *Molec. cell. Biol.* **8**, 2933-2941 (1988).
16. Dionne, C. A. et al. *EMBO J.* **9**, 2685-2692 (1990).
17. Mansukhani, A., Moscatelli, D., Talarico, D., Levytska, V. & Basilico, C. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4378-4382 (1990).
18. Orr-Urtreger, A., Givol, D., Yayon, A., Yarden, Y. & Lonai, P. *Development* **113**, 1419-1434 (1991).
19. Peters, K. G., Werner, S. & Williams, L. T. *Development* **114**, 233-243 (1992).
20. Morgan, B. A., Izpisua-Belmonte, J.-C., Duboule, D. & Tabin, C. J. *Nature* **358**, 236-239 (1992).

21. Dollé, P., Izpisua-Belmonte, J.-C., Falkenstein, H., Renucci, A. & Duboule, D. *Nature* **342**, 767-772 (1989).
22. Takawa, Y., Ohse, C., Wang, E. A., Wozney, J. M. & Yamashita, K. *Biochem. biophys. Res. Commun.* **174**, 96-101 (1991).
23. Wilkinson, D. G., Bhatt, S. & McMahon, A. P. *Development* **105**, 131-136 (1989).
24. Haub, O. & Goldfarb, M. *Development* **112**, 397-406 (1991).
25. Smith, D. B. & Johnson, K. S. *Gene* **67**, 31-40 (1988).
26. Hébert, J. M., Basilico, C., Goldfarb, M., Haub, O. & Martin, G. R. *Dev. Biol.* **138**, 454-463 (1990).
27. Wanek, N., Muneeoka, K., Holler-Dinsmore, G., Burton, R. & Bryant, S. V. *J. exp. Zool.* **249**, 41-49 (1989).
28. Neubert, D. & Barrach, H.-J. in *Methods in Prenatal Toxicology* (eds Neubert, D., Merker, H.-J. & Kwasigroch, T. E.) 241-251 (Georg Thieme, Stuttgart, 1977).
29. Frohman, M. A., Boyle, M. & Martin, G. R. *Development* **110**, 589-607 (1990).
30. Dollé, P., Izpisua-Belmonte, J.-C., Boncinelli, E. & Duboule, D. *Mech. Dev.* **36**, 3-13 (1991).
31. Aono, H. & Ide, H. *Dev. Biol.* **128**, 136-141 (1988).
32. Hinchliffe, J. R. & Johnson, D. R. *The Development of the Vertebrate Limb: An Approach through Experiment, Genetics and Evolution* (Clarendon, Oxford, 1980).

ACKNOWLEDGEMENTS. We thank J. Fallon, C. Tickle, G. Morriss-Kay, H. Haack, P. Martin and members of our laboratory for discussions, J. de la Torre for advice on the BrdU studies, and C. Murphy for help with preparation of the manuscript. This work was supported by NIH grants to G.R.M. L.N. is supported by a fellowship from the American Cancer Society.

A gene for the mouse pink-eyed dilution locus and for human type II oculocutaneous albinism

Eugene M. Rinchik*, Scott J. Bultman*†, Bernhard Horsthemke‡, Seung-Taek Lee§, Kathleen M. Strunk§, Richard A. Spritz§, Karen M. Avidano||, Michelle T. C. Jong|| & Robert D. Nicholls||¶

* Biology Division and †University of Tennessee–Oak Ridge Graduate School of Biomedical Sciences, Oak Ridge National Laboratory, PO Box 2009, Oak Ridge, Tennessee 37831, USA

‡ Institut für Humangenetik, Universitätsklinikum Essen, Hufelandstrasse 55, D-4300 Essen 1, Germany

§ Department of Medical Genetics, University of Wisconsin, Madison, Wisconsin 53706, USA

|| Department of Neuroscience, Box 100244 JHMC, University of Florida College of Medicine, Gainesville, Florida 32610, USA

¶ To whom correspondence should be addressed

THE mouse pink-eyed dilution (*p*) locus on chromosome 7 is associated with defects of skin, eye and coat pigmentation¹. Mutations at *p* cause a reduction of eumelanin (black–brown) pigment and altered morphology of black pigment granules (eumelanosomes), but have little effect on pheomelanin (yellow–red) pigment². We show here that the human complementary DNA *DN10*, linked to the *p* locus in mice^{3–5}, identifies the human homologue (*P*) of the mouse *p* gene, and appears to encode an integral membrane transporter protein. The expression pattern of this gene in various *p* mutant mice correlates with the pigmentation phenotype; moreover, an abnormally sized messenger RNA is detected in one mutant, *p*^{un}, which reverts to the normal size in *p*^{un} revertants. The human *P* gene corresponds to the *D15S12* locus within the chromosome segment 15q11–q13, which is typically deleted in patients with Prader–Willi and Angelman syndrome (see ref. 5 for review). These disorders are phenotypically distinct, depending on the parent of origin of the deleted chromosome^{5–7}, but both syndromes are often associated with hypopigmentation of the skin, hair and eyes (see ref. 8 for review), and deletion of the *P* gene may be responsible for this hypopigmentation. In addition, we report a mutation in both copies of the human *P* gene in one case of tyrosinase-positive (type II) oculocutaneous albinism, recently linked to 15q11–q13 (ref. 9).

To extend previous recombinant-inbred strain analyses demonstrating linkage of *D15S12h* (formerly designated *D7Nic1*) to the mouse *p* locus^{3,4}, we performed an interspecific backcross analysis. No recombinants were detected between *p* and *D15S12h* in 182 segregants from the cross (129/Rl-*pc*^{ch}/*pc*^{ch} × *M. spretus*) F₁ × 129/Rl-*pc*^{ch}/*pc*^{ch} (Fig. 1a), implying a map distance of less than 1 centimorgan (two-sided 95% confidence limits, 0–2 cM). The complete linkage of *D15S12h* and *p* thus suggested that *D15S12h* might be deleted in radiation-induced *p*-locus mutations, and Southern blot hybridization analysis demonstrated that all DNA fragments hybridizing to *DN10* are deleted in the mouse mutations *p*^{7FR6OLb}, *p*^{6H}, *p*^{46DFIOD}, *p*^{116G} and *p*^{25DVT} (Fig. 1b).

The *DN10* cDNA detects an RNA expression pattern in mouse tissues consistent with the phenotypic effects of *p* mutations. *DN10* recognizes two transcripts of 3.4 kilobases (kb) and 1.9 kb (Fig. 2a). The 3.4-kb mRNA, but not the 1.9-kb mRNA, derives from the *D15S12h* locus, because it is not present in skin of *p*^{6H}/*p*^{6H} mice (Fig. 2a, lane 14), in which *D15S12h* is deleted (Fig. 1b). (Thus the 1.9-kb mRNA is derived from a *DN10*-related locus outside the *p*^{6H} deletion (see also Fig. 2 legend) and will not be considered further.) The 3.4-kb mRNA is expressed highly in B16 melanoma, in albino (*c/c*) skin (which contains melanocytes but no pigment) and in skin

FIG. 1 Molecular analysis of the *D15S12h* locus in mice. a, Segregation of *D15S12h* haplotypes in 182 members of an interspecific backcross²² (129/Rl-*pc*^{ch}/*pc*^{ch} × *M. spretus*) × 129/Rl-*pc*^{ch}/*pc*^{ch}. Filled boxes represent 129/Rl *Bam*HI restriction fragment length variants (RFLVs) (6 kb and 3.3 kb); open boxes represent *M. spretus* RFLVs (7.7 kb and 3.6 kb). Numbers of animals in each class are indicated. b, Southern blot analysis of lethal *p*-locus mutations with the human *DN10* cDNA probe. DNA from mice of the indicated genotypes was digested with *Bam*HI and hybridized to *DN10* (upper panel) or control probe 23.3 from the more distal chromosome-7 *Emv-23* locus²³ (lower panel). Designations are abbreviated forms of the complete genotype (for example, 6H represents *p*^{6H}). c, Southern blot analysis of *p*^{un}/*p*^{un} and control DNA. *Hind*III-digested DNA (lanes 1–4) from *p*^{un}/*p*^{un}, *p*^{6H}/*p*^{6H}, and inbred strains C57BL/6J and AKR/J were hybridized to *DN10*. Arrows indicate fragments duplicated in *p*^{un} DNA (lane 1); all other bands are of equal intensity, with the exception of the 1.6-kb fragment, which is too faint to determine. Bands detected by subfragments of *DN10* (*Eco*RI–*Taq*I 912, *Taq*I–*Taq*I 1412; see Fig. 4 for designation) are shown to the right. Our data indicate the duplication is >18.6 kb; re-evaluation of the data in ref. 12 shows that the minimum size of the duplication is 51.8 kb. METHODS. Mice carrying *p*^{6H}, and DNA from *p*^{6H}/*p*^{6H}, *p*^{un}/*p*^{un}, and inbred strains were purchased from the Jackson Laboratory. Breeding stocks carrying *p*^{un} were provided by H. G. Wolfe, University of Kansas. All other mice were bred at Oak Ridge National Laboratory. Lethal mutations at the chromosome-7 pink-eyed dilution (*p*) locus were recovered²⁴ opposite the original *p* mutation, and are maintained by outcross to +/+ mice, with progeny testing for *p*. Neonates or juvenile mice homozygous for *p*^{7FR6OLb} or *p*^{6H}, respectively, or compound heterozygotes involving *p*^{6H} or *p*^{7FR6OLb} and any of the prenatally lethal mutations *p*^{46DFIOD}, *p*^{116G}, or *p*^{25DVT}, were identified in the progeny of a cross of appropriate heterozygotes by their pink eye and/or dilute fur colour. The *p*^{un+1Rn} reversion of *p*^{un} was recognized in a mottled pink-eyed-dilute//wild-type mosaic. Wild-type heterozygotes (*p*^{un}/*p*^{un+1Rn}) were obtained from a cross of the mosaic by *p*^{un}/*p*^{un}. Wild-type homozygotes (*p*^{un+1Rn}/*p*^{un+1Rn}), obtained from a cross of heterozygotes, were identified by progeny tests to *p/p* mice. *DN10* is a 3.2-kb cDNA from a human fetal brain library³ that detects 12 fragments totalling 110 kb in *Hind*III-digested human DNA; all fragments but one (see Fig. 3 legend) map to 15q11–q13 on the basis of hybridization to DNA from somatic cell hybrids (unpublished data). A human *MYOD1* cDNA²² and the pIGF-1R-8 cDNA²⁵ were used as probes for the *Myod-1* and *Igf1r* loci, respectively. DNA preparation and Southern blot hybridization were as described^{18,23}, except that mouse filters hybridized to *DN10* received a final wash of either 1 × SSC at 65 °C or 0.1 × SSC at 39 °C.

from agouti (*A/A*; combination of yellow and black) and non-agouti (*a/a*; black) mice (Fig. 2a, lanes 3, 4, 9, and 10). The 3.4-kb mRNA is not expressed in *W/W*^v skin (lane 17), which lacks melanocytes, or in any skin that exclusively produces yellow pigment due to mutations at the agouti locus (lanes 6, 7, and 8). Of particular interest is the expression in *a^l/a^l* mice, in which the 3.4-kb mRNA is expressed in the black dorsal skin but not in the yellow ventral skin (lanes 5 and 6), in contrast to expression in black ventral skin of *a/a* mice (lane 10). Thus in cases in which pigment is produced, the presence of the 3.4-kb transcript correlates completely with the presence of black pigment, consistent with the effects of *p* mutations on eumelanosomes, while having little or no effect on pheomelanosomes^{1,2}.

The 3.4-kb mRNA is expressed in skin of mice homozygous or hemizygous for the original *p* mutation, but at a very reduced level (Fig. 2a, lanes 11 and 12), consistent with genetic data¹⁰ indicating that *p* is a low activity rather than a null mutation. Of particular importance is the result obtained in *p*^{un}/*p*^{un} skin. The *p*^{un} (pink-eyed unstable) mutation results in a phenotype indistinguishable from that of *p/p* homozygotes, except that *p*^{un} reverts to wild type at a high frequency¹¹ and is associated with genomic duplication¹². The *DN10* probe recognizes an aberrant 4.3-kb mRNA in *p*^{un}/*p*^{un} skin (Fig. 2a, lane 1), but detects a correctly sized 3.4-kb mRNA in the normally pigmented skin of a homozygous *p*^{un+1Rn}/*p*^{un+1Rn} revertant (Fig. 2a, lane 2). This is consistent with our observation that 5' and 3' *DN10* probes (912 and 1412, respectively) detect both duplicated and non-duplicated restriction fragments in Southern blot analyses

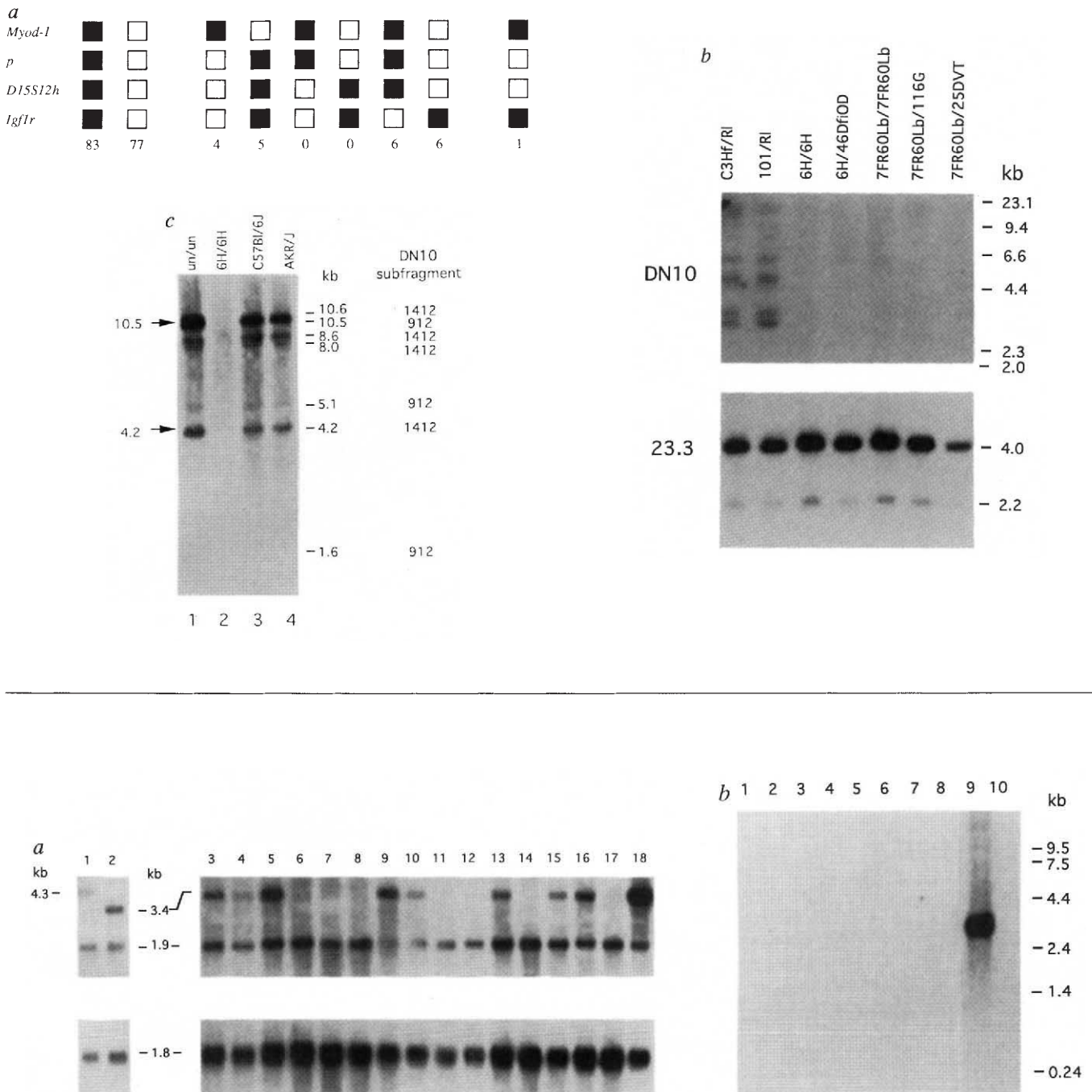


FIG. 2 Northern blot analysis of mouse and human tissue RNA with the *DN10* cDNA. **a**, Mouse blot using (upper panel) *DN10* probe or (lower panel) chicken tubulin control probe. Sizes of the transcripts (in kb) are given. Poly(A)⁺ RNA is from mouse skin (lanes 1–17) or from B16 melanoma (lane 18). Lane 1, *a/a*; *p^{un}/p^{un}*, postnatal d8, whole-trunk, non-agouti pink-eyed dilute; lane 2, *a/a*; *p^{un+1Rn}/p^{un+1Rn}*, d8, whole-trunk, non-agouti black; lane 3, *+/+*, d11, dorsal, agouti; lane 4, *+/+*, d11, ventral, agouti; lane 5, *aⁱ/aⁱ*, d9, dorsal, black; lane 6, *aⁱ/aⁱ*, d9, ventral, yellow; lane 7, *A^y/+*, d6, dorsal, yellow; lane 8, *A^y/+*, d6, ventral, yellow; lane 9, *a/a*, d6, dorsal, non-agouti black; lane 10, *a/a*, d6, ventral, non-agouti black; lane 11, *a/a*; *p/p*, d7, whole-trunk, non-agouti pink-eyed dilute; lane 12, *p/p*, d4, whole-trunk, agouti-pink-eyed dilute; lane 13, *p/+*, d4, whole-trunk, agouti; lane 14, *a/a*; *p^{6H}/p^{6H}*, d7, 10, whole-trunk, non-agouti pink-eyed dilute; lane 15, *a/a*; *p^{6H}/p^{7R75M}*, d10, whole-trunk, non-agouti dark pink-eyed dilute; lane 16, *c/c*, d6, whole-trunk, albino, no pigment; lane 17, *W/W^v*, d6, whole-trunk, albino, no melanocytes; lane 18, *a/a*, B16 melanoma, non-agouti black. *p^{7R75M}* is a viable, radiation-induced mutation of *p* that has minor effects on pigmentation. See ref. 1 for a review of genotypes and phenotypes. **b**, Poly(A)⁺ RNA was used from the following human tissues or cells: pancreas (lane 1),

kidney (lane 2), skeletal muscle (lane 3), liver (lane 4), lung (lane 5), placenta (lane 6), brain (lane 7), heart (lane 8), melanocytes (lane 9), and HEL erythroleukaemia cells (lane 10).

METHODS. Poly(A)⁺ RNA was prepared by standard procedures²⁶ and analysed by electrophoresis through 1% agarose/0.22 M formaldehyde (mouse) or 1% agarose/0.92 M formaldehyde (human melanocytes and HEL cells) and hybridization to radiolabelled probes. A northern blot containing 2 µg each of polyadenylated mRNAs from various human tissues was purchased from Clontech. The hybridization probe was prepared by polymerase chain reaction (PCR) using oligonucleotides RN19 (5'-CCATATCCATGCTGTTCTGC-AATCCC-3') and RN48 (5'-CCCAGCCTGACCCATGTGGTGG-3') as primers and plasmid containing *DN10* as template, followed by gel purification of the 1.4-kb PCR product. After hybridization, the final wash was at 60 °C in 0.1 × SSPE and 0.5% SDS. The 1.9-kb mRNA in the mouse (see text), detected by both *DN10* and a 2.1-kb probe derived from the *DN10* cDNA by PCR using primers RN14 (5'-GGATCAAGGAAAGCTCTGGCAGCTG-3') and RN19, is present in all tissues examined (data not shown). Probes from the co-ligated sequence 3' to the poly(A) recognition site (Fig. 3) in *DN10* did not hybridize to mouse RNA.

1561 GGACITTTGCCGGATTCACTGCACACATGTTTCATTGGGATTTCGGCTTCTCTGGTCTG 523
D F A G F T A H M F I G I C L V L L V C

1621 CTTTCCGCTCCTCAGACTCCTTTACTGGAACAGAAAGCTTTATAACAGGAACCCAGTGA 543
F P L L R L L Y W N R R K L Y N K E P S E

1681 GATTGTTGAACTGAAGCACGAGATTCACTGCTGGCGCCTGACTGCTCAGCGCATCAGCCC 563
I V E L K H E I H V W R L T A Q R I S P

1741 GGCCAGCCGCGAGGAGACAGCTGTGGCGCGCTGCTGCTGGGGAAGGTGCTGGCACTGGA 583
A S R E E T A V R R L L L G K V L A L E

1801 GCACCTGCTCGCCCGGAGGCTGCACACCTTCCACAGACAGATCTCACAGGAGACA AAAA 603
H L L A R R L H T F H R Q I S Q E D K N

1861 TTGGGAGACCAATATCCAAGAACTCCAAAAAGCATAGGATATCTGACGGGATTCTGCT 623
W E T N I Q E L Q K K H R I S D G I L L

1921 CGCCAAATGCCTGACAGTGTGGGATTGTTATCTTCATGTTTTTCTCAATTCTGTTTGT 643
A K C L T V L G G F V I F M F F L N S T F V

1981 CCTTGGCATTCATCTTGATCTGGGATGCTATTCTGGGTCCCATCTGGTTCGTAAT 663
P G I H L D L G W I A I L G A I W L L I

2041 TTTAGCTGATATTCATGATTTTGAGATAATTTACACAGAGTGAATGGGCAACCCITCT 683
L A D I H D F E I I L H R V E W A T L L

2101 GTTTTTCGAGCGCTCTTTGTTCTGATGGAGGATTGGCCATTCACCTTAATAGAATA 703
F F A A L F V L M E A L A H L H L I E Y

2161 TGTGGAGAACAACTGCTTTTGCTAATAAAGATGGTCCACAGGAGCAGCGCCTCATAGC 723
V G E Q T A L L I K M V P E E Q R L I A

2221 CGCCATTGTCTCGTGGTGTGGGTCTCAGCGCTGGCGCTGCTCCCTGATTGACAACATCCC 743
A I V L V V W V S A L A S S L I D N I P

2281 GTTCACTGTACCATATTCCTGCTCCTGAACCTGAGCCACGACCTGAGGTTGGCCT 763
F T A T M I P V L L N L S H D P V G L

2341 GCGCGACCGCGCTCATGTATGCCCTGGCCTTCGGTCTTGCCTGGGAGGCAACGGGAC 783
F A P P L M Y A L A F G A C L G G N G T

2401 ACTGATTGGCGCGTCGGCAACCTCGTGTGTGCAGGATTGCAGAACAGCATGGATATGG 803
L T I C A S A N V V C A G I A E Q H G Y G

2461 GTTCTCCTTACGAAATTTTCAGGCTGGGCTTCCCAATGATGGTTTGTTCTTCGACTGT 823
F S F M E F F R L G F P M M V V S C T V

2521 TGGGATGTTTATCTCTTGTGGCTCATGTGGTGGTGGGATGGAATTAATAGACATCCAT 843
G M C Y L L V A H V V V G W N * *

2581 CTATTGCTCGAAGACTAAAGGAAACTTCATCCATCACAACCCATTAGTCATAAACTACC 863

2641 CTGACCCCACTGTTTGAAGAAGAAAAGTGCTTACCCTGGAGATGCTACAGACACAGT 883

2701 GGAATAGACCTTGACACTAACACTCTAATTCAAGCGAATGTGGAAACACCATGACCTCT 903

2761 CTGTGTGTCTCTTCTCCCCAAGGACAAAATGTAGAAGATGTGAGATAACTTTACTCAAG 923

2821 TTCCCCTCCAGAAAAATACGTATGTTTAAAAACCCCTCTCTATATACATAGGAAAAGACA 943

2881 CACATCCACCTAAAAATTGACTGTACTGTTAATCTGCAATTCTCCTGAGGCTAAACACAG 963

2941 TTTGTTTTTCTGTGAATCACTTTTCACTGTTAAAATAATCAGCAATCAAAATGTATGCTTT 983

3001 CTGAATATAGACTTTCTGGGAAAAGGTTTACTGCTCGTAAGGAAACATTTTATGTATTAA 1003

3061 AATAAACTGT

b

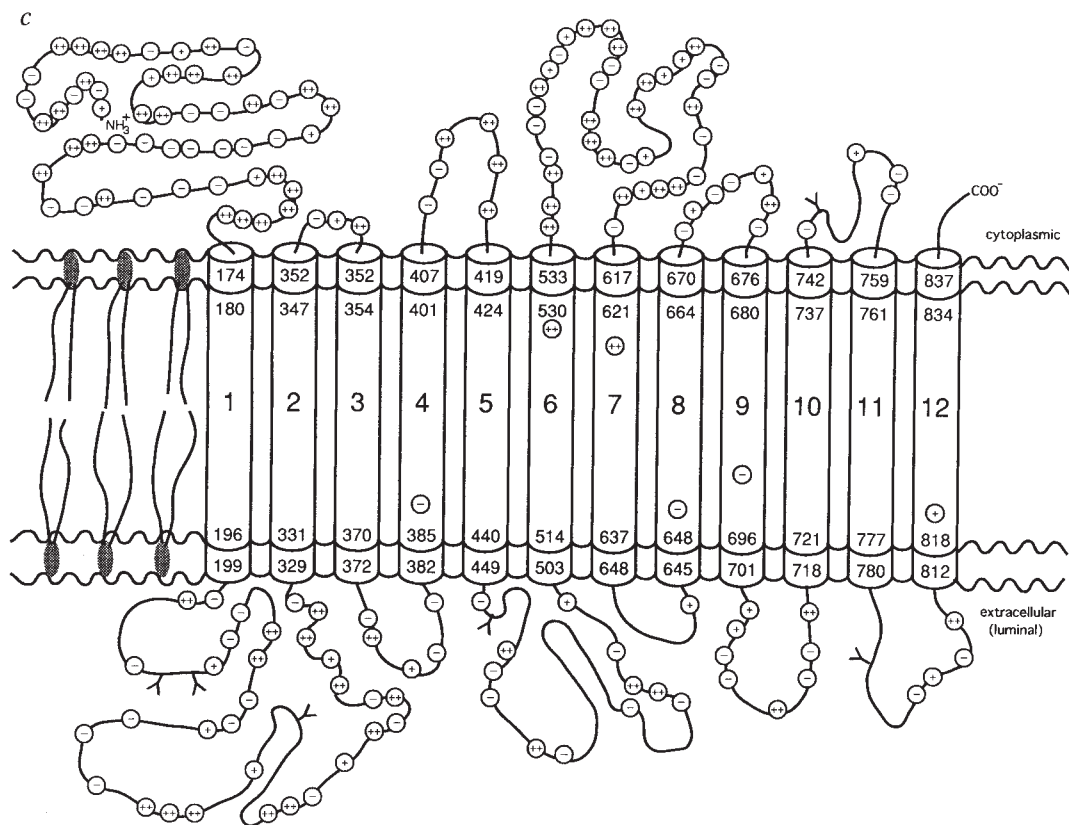
FIG. 3 Molecular characterization of human *P* cDNA and the putative P polypeptide. *a*, Nucleotide and deduced amino-acid sequences of human *P* cDNA. Composite DNA sequences of *DN10* plus three additional 5' *P* cDNAs are shown. Twelve putative transmembrane domains^{27,28} are boxed; potential *N*-glycosylation sites are indicated by bars, potential protein kinase C phosphorylation sites by triangles, and translational termination codons by asterisks; the polyadenylation signal is double-underlined and a melanocyte-specific downstream element²⁹ is singly underlined. The cDNA sequence presented ends where the genomic sequence (unpublished data) diverges; the *DN10* cDNA clone contains another 547 nucleotides that represent a co-ligated cDNA derived from an unrelated non-chromosome-15 locus (data not shown). The human *P* cDNA sequence has been deposited in GenBank (accession number M99564). *b*, DNA sequence of a 77-bp minor alternatively spliced exon of the human *P* gene. A partial nucleotide sequence of genomic clone¹⁸ *IR10-1* is illustrated, aligned with the *DN10* cDNA and predicted amino-acid sequences.

of p^{un}/p^{un} DNA (Fig. 1c shows results with the complete *DN10* cDNA), but not in the revertant (data not shown), implying that the duplication is intragenic. These expression and structural data for p^{un} provide strong evidence that the *DN10* cDNA corresponds to the *p* gene itself, because reversion to wild-type pigmentation is accompanied by a restoration of the normal *DN10/D15S12h* transcript. We designate the human homologue of the murine *p* locus by the gene symbol *P*.

A probe derived from *DN10* hybridized to an abundant 3.4-kb transcript in poly(A)⁺ RNA from normal primary human melanocytes (Fig. 2b), which is not present in human pancreas, kidney, skeletal muscle, liver, lung, placenta, heart, HEL erythroleukaemia cells or adult brain (Fig. 2b). The DNA sequence and deduced amino-acid sequence of the human *P* polypeptide was determined from four cDNA clones (Fig. 3a). The *DN10* cDNA clone contains an alternatively spliced exon present in only a small fraction of *P* mRNA from human melanocytes and fetal brain (Fig. 3b). The sequence of the putative human *P*

polypeptide predicts an integral membrane protein with 12 transmembrane domains (Fig. 3a and c). Searches of nucleic acid (GenBank/EMBL rel. 71) and protein (NBRF-PIR rel. 32) databases showed the best match (21% identity) to a tyrosine-specific transport protein from *Escherichia coli*¹³ (accession number, GRECY), which appears to contain 12 transmembrane domains. A structural arrangement of 6–12 transmembrane domains is common among mammalian proteins that transport small organic molecules¹⁴. We speculate that the *P* polypeptide may be a component of the melanosomal membrane, possibly involved in the transport of tyrosine, the primary precursor to melanin biosynthesis¹⁵. In this regard, it is intriguing that pigmentation of melanocytes from *p/p* mice¹⁶, and also from human patients with type II oculocutaneous albinism (OCA)¹⁷, is greatly increased by incubation in excess exogenous tyrosine.

Deletion of *P* might be responsible for the mild hypopigmentation usually found in Prader-Willi (PWS) and Angelman (AS) patients with 15q11–q13 deletions (see ref. 8 for review).



Inclusion of the *IR10-1* exon introduces a frameshift and translational termination at codon 669 of the minor *P*-derived mRNA. c, Model for arrangement of the *P* polypeptide in the membrane. Charged amino acids are shown as circles (negative charge D, E; +, H; ++, K, R), with neutral amino acids in the connecting line. Putative transmembrane segments are depicted as cylinders 1–12; numbers indicate the inner and outer boundary^{28,30}. Six potential *N*-glycosylation sites are denoted by forks.

METHODS. *DN10* cDNA extends from nucleotides 358–3,070. Three additional human *P* cDNAs were obtained by screening two human melanocyte cDNA libraries using 370-bp (nt 358–735) and 260-bp (nt 118–358) fragments from the 5' end of *P* cDNAs. Poly(A)⁺ mRNA was prepared from 5×10^6 log phase normal primary human melanocytes (NHEM337; Clonetics) and 1 μ g used to prepare oligo(dT)-primed cDNA as described²⁶. One-twentieth of the reaction was used as template for 30 cycles of PCR with different pairs of *DN10*-derived primers. Parallel PCRs were carried out using 0.1 μ g normal human genomic DNA as template. One-fifth of each PCR was analysed by electrophoresis in a 5% polyacrylamide gel. Most *DN10* primer pairs did not amplify any products when normal human genomic

DNA was used as template (data not shown), suggesting that the human *P* gene contains many large intervening sequences. In contrast, all but one primer pair amplified the expected size fragments from melanocyte cDNA. Primer pair RN46 (5'-CCCTGGCATTTCATCTTGATC-3')/RN50 (5'-CTATTAAGTGGAGATGTGCC-3') amplified a major 174-bp product from melanocyte cDNA, 77 bp shorter than the expected 251-bp product. This pair of primers spans a segment corresponding to an exon present in genomic clone¹⁸ *IR10-1*, which was the probe originally used to isolate the *DN10* cDNA³. To determine whether the 174-bp RN46/RN50 PCR amplification product corresponds to cDNA lacking the *IR10-1* exon, we cloned this PCR fragment in M13mp18 and determined the DNA sequence of six clones; the *IR10-1* exon was precisely deleted from each, but otherwise agreed with the nucleotide sequence of *DN10* (data not shown). Similar results were obtained from a human melanocyte cDNA library and human fetal brain cDNA (data not shown). Thus most of the *P* mRNA in human melanocytes and in fetal brain lacks the *IR10-1* exon; the sequence of this predominant transcript is shown in a. DNA sequences were analysed using the DNASTAR and PC/GENE software packages and IALIGN³¹.

In fact, normal pigmentation is observed in a PWS patient with an unbalanced translocation in which all loci except *D15S12* are deleted¹⁸.

A surprisingly large number of PWS and AS patients also exhibit clinical features characteristic of tyrosinase-positive (type II) OCA, including severe hypopigmentations, decreased visual acuity, nystagmus and strabismus^{19–21} (J. Clayton-Smith, S. Saitoh, personal communications, and our unpublished data); these atypical PWS and AS patients may be hemizygous for mutant alleles of a gene responsible for type II OCA. Accordingly, we carried out Southern blot hybridization analysis of the *P* gene in a PWS patient who also has type II OCA. Karyotype analysis demonstrated a *de novo* deletion of 15q11–q13, and molecular analysis demonstrated that the deletion included *D15S12* and other proximal markers typically deleted in PWS^{5,8,21}. In contrast to 15 other PWS/AS patients who do not have OCA (unpublished data), probes from both the 3' (Fig. 4, right) and central (Fig. 4, left) part of the *DN10* cDNA identify homozygous deletion of a 6.6-kb genomic *Hind*III fragment in the PWS/OCA patient. Furthermore, this fragment was of reduced intensity in the mother (Fig. 4). All other bands detected by the cDNA (with one exception: see Fig. 4 legend) are normal, except for the expected reduced intensity in the patient. These data demonstrate that the patient has a partial deletion of the *P* locus in the chromosome 15 inherited from his mother, and that the entire locus is absent from the paternally derived chromosome. The prevalence of type II OCA among PWS and AS patients, perhaps 1% (refs 19–21; J. Clayton-Smith, S. Saitoh, personal communications, and our unpublished data), is consistent with the expected frequency of carriers of type II OCA, given a frequency of the disease of ~1 per 36,000 in caucasians¹⁷. Recently, genetic linkage has been demonstrated⁹ between type II OCA and markers in 15q11–q13. Together these findings indicate that mutations of the human *P* gene may account for a significant fraction of cases of type II OCA, the most frequent form of oculocutaneous albinism worldwide¹⁷.

The pigmentation defects observed in PWS and AS constitute one component of these contiguous gene syndromes, independent of the primary, imprinted genetic element(s) that result in the major clinical features of these syndromes^{5,8}. We have shown here that specific defects in this single gene (*P*) are also likely to result in the phenotype of type II oculocutaneous albinism in man and pink-eyed dilution (*p*) in mouse. This study

exemplifies the combined use of human and mouse genetics to dissect human genetic diseases involving multiple genes and complex phenotypes.

Note added in proof: While this letter was under review, J. M. Gardner *et al.*³² also reported identification of the mouse *p* gene. □

Received 28 August; accepted 10 November 1992.

1. Silvers, W. K. *The Coat Colors of Mice. A Model for Mammalian Gene Action and Interaction* (Springer, New York, 1979).
2. Russell, E. S. *Genetics* **34**, 146–166 (1949).
3. Nicholls, R. D. *Mouse Newsletter* **84**, 87–88 (1989).
4. Nakatsu, Y., Gondo, Y. & Brilliant, M. H. *Mammalian Genome* **2**, 69–71 (1992).
5. Nicholls, R. D., Rinchik, E. M. & Driscoll, D. J. in *Genomic Imprinting in Mouse and Man* (ed. Surani M. A. & Reik, W.) *Seminars dev. Biol.* **3**, 139–152 (1992).
6. Cassidy, S. B. *Curr. Problems Pediatr.* **14**, 1–55 (1984).
7. Clayton-Smith, J. & Pembrey, M. E. *J. med. Genet.* **29**, 412–415 (1992).
8. Nicholls, R. D. *Am. J. med. Genet.* (in the press).
9. Ramsay, M. *et al.* *Am. J. hum. Genet.* **51**, 879–884 (1992).
10. Russell, L. B. *Prog. Top. Cytogenet.* **3**, 205–250 (1983).
11. Melvold, R. W. *Mutation Res.* **12**, 171–174 (1971).
12. Brilliant, M. H., Gondo, Y. & Eichler, E. M. *Science* **252**, 566–569 (1991).
13. Wookey, P. J. & Pittard, A. J. *J. Bact.* **170**, 4946–4949 (1988).
14. Wang, H., Kavanaugh, M. P., North, R. A. & Kabal, D. *Nature* **352**, 729–731 (1991).
15. Hearing, V. J. & Tsukamoto, K. *FASEB J.* **5**, 2902–2909 (1991).
16. Sidman, R. L. & Pearlstein, R. *Dev. Biol.* **12**, 93–116 (1965).
17. Witkop, C. J. *et al.* in *The Metabolic Basis of Inherited Disease* Vol. II, 6th edn (ed. Scriver, C. R. *et al.*) 2905–2947 (McGraw-Hill, New York, 1989).
18. Tantravahi, U. *et al.* *Am. J. med. Genet.* **33**, 78–87 (1989).
19. Fryburg, J. S., Breg, W. R. & Lindgren, F. *Am. J. med. Genet.* **38**, 58–64 (1991).
20. Wallis, C. E. & Beighton, P. H. *J. med. Genet.* **26**, 337–339 (1989).
21. Daumer, C. & Schuffenhauer, S. *Med. Genetik* **3**, 27–29 (1991).
22. Scrabble, H. J., Johnson, D. K., Rinchik, E. M. & Cavenee, W. K. *Proc. natn. Acad. Sci. U.S.S.* **87**, 2182–2186 (1990).
23. Rinchik, E. M., Machanoff, R., Cummings, C. C. & Johnson, D. K. *Genomics* **4**, 251–258 (1989).
24. Rinchik, E. M. & Russell, L. B. in *Genome Analysis Vol I* (eds Davies, K. & Tilghman, S.) 121–158 (Cold Spring Harbor Laboratory Press, New York, 1990).
25. Ullrich, A. *et al.* *EMBO J.* **5**, 2503–2512 (1986).
26. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* 2nd edn (Cold Spring Harbor Laboratory Press, New York, 1989).
27. Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105–132 (1982).
28. Klein, P., Kanehisa, M. & Delisi, C. *Biochim. biophys. Acta* **815**, 468–476 (1985).
29. Shibahara, S. *et al.* *J. biol. Chem.* **266**, 15895–15901 (1991).
30. Lodish, H. F. *Trends Biochem. Sci.* **13**, 332–334 (1988).
31. Dayhoff, M. O., Barker, W. C. & Hunt, L. T. *Meth. Enzym.* **91**, 524–545 (1983).
32. Gardner, J. M. *et al.* *Science* **257**, 1121–1124 (1992).

ACKNOWLEDGEMENTS. We thank R. Neve, P. Rogan, H. G. Wolfe, C. Daumer, S. Schuffenhauer, L. Taylor, K. Houser, L. Biddle, W. Gottlieb, A. Garcia, L. Stubbs, R. Woychik, L. Russell, M. Kilberg and G. Shaw. This work was supported by the Office of Health and Environmental Research, US Department of Energy under contract with Martin Marietta Energy Systems Inc. (EMR), the Deutsche Forschungsgemeinschaft (B.H.), the NIH Health (R.A.S.), the March of Dimes Birth Defects Foundation (R.A.S. and R.D.N.), the American Heart Foundation, Florida Affiliate (R.D.N.), and the Pew Scholars Program in Biomedical Sciences (R.D.N.).

Modulation of the cGMP-gated channel of rod photoreceptor cells by calmodulin

Yi-Te Hsu & Robert S. Molday*

Department of Biochemistry, Faculty of Medicine, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada

PHOTOLEACHING of rhodopsin in rod photoreceptors activates the visual cascade system leading to a decrease in cyclic GMP and the closure of cGMP-gated channels in the rod outer segment plasma membrane^{1–4}. Calcium plays an important role in the recovery of the rod outer segment to its dark state by regulating the resynthesis of cGMP by guanylate cyclase^{5–7}. Here we report that calmodulin, a Ca²⁺-binding protein present in the rod outer segment^{8,9}, increases the apparent Michaelis constant of the channel for cGMP. This results in a decrease in the rate of cation influx into the rod outer segment by two- to sixfold at low cGMP concentrations and has the effect of increasing the sensitivity of the channel to small changes in cGMP levels during phototrans-

* To whom correspondence should be addressed.

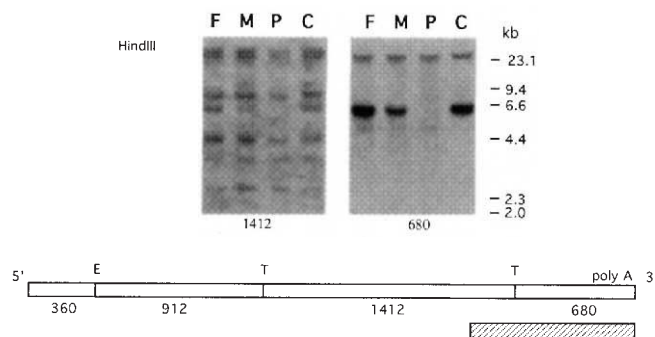


FIG. 4 An inherited *P* (*D15S12*)-locus deletion in a patient with PWS and tyrosinase-positive oculocutaneous albinism. *Hind*III-digested DNA from the patient (P), his father (F), mother (M), and a control (C) were hybridized to the 1,412-bp (left panel) or 680-bp (right panel) probes derived from the *DN10* cDNA. Similar results were obtained with *Eco*RI, *Bam*HI and *Kpn*I (data not shown). Hatched box illustrates the position of the inherited deletion demonstrated in the patient and mother; open boxes, cDNA clones; E, *Eco*RI; T, *Taq*I; poly A, polyadenylation site. The deletion detected by the 1,412-bp probe includes only the last exon of the *P* gene (Fig. 3a, and unpublished data). The second fragment (23 kb) detected by the 680-bp probe is from an unrelated locus (see Fig. 3 legend) and serves as a control for the amount of DNA loaded in this experiment as it is not deleted. A clinical and cytogenetic description of the proband with PWS and type II OCA is given in ref. 21.

duction. Biochemical studies indicate that calcium-calmodulin binds to a protein of M_r 240K which is tightly associated with the channel¹⁰. On the basis of these studies, Ca^{2+} is suggested to play a central role in photorecovery and light adaptation, not only by regulating guanylate cyclase, possibly through recoverin^{6,7}, but also by modulating the cGMP-gated channel through calmodulin interaction with the 240K protein.

The effect of calmodulin on the cGMP-gated channel activity was investigated using a calcium ion flux assay. The Ca^{2+} -sensitive dye, Arsenazo III, was trapped inside the rod outer segment (ROS) membrane vesicles and the rate of Ca^{2+} influx was measured spectrophotometrically¹¹ as a function of cGMP concentration in the presence and absence of calmodulin (Fig. 1a, b). The apparent Michaelis constant (K_m) for cGMP increased from $19 \pm 0.4 \mu\text{M}$ in the absence of calmodulin to $33 \pm 2 \mu\text{M}$ in the presence of Ca^{2+} -calmodulin. This change in K_m translates into a relatively large change in the rate of Ca^{2+} influx (sixfold) at $12.5 \mu\text{M}$ cGMP. The V_{max} and the cooperativity for cGMP as determined by a Hill plot, however, was unaffected by Ca^{2+} -calmodulin. Other calcium-binding proteins, including recoverin and S-100 protein, had no effect on channel activity under similar conditions. The Ca^{2+} -calmodulin effect on channel activity was also inhibited by excess mastoparan, a calmodulin inhibitor¹².

A neutral red dye assay¹³ was used to determine the effect of Ca^{2+} -calmodulin on the cGMP-dependent influx rate of different cations into ROS membrane vesicles. At $20 \mu\text{M}$ cGMP, Ca^{2+} -calmodulin decreased the initial rate of ion influx by 2–3-fold for all cations except Mg^{2+} (Fig. 2). The smaller decrease for Mg^{2+} (1.4-fold) may be due to the interference of high Mg^{2+} concentration on the interaction of Ca^{2+} with calmodulin¹⁴. For Ca^{2+} influx rates using the neutral red assay,

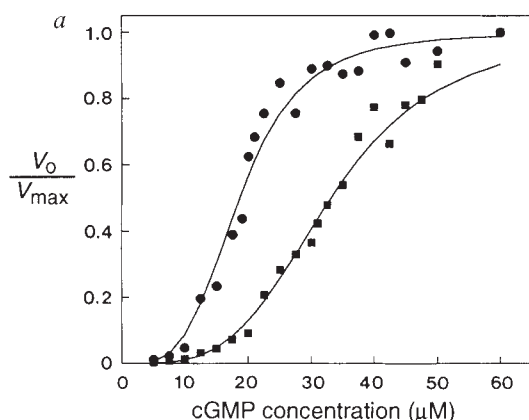


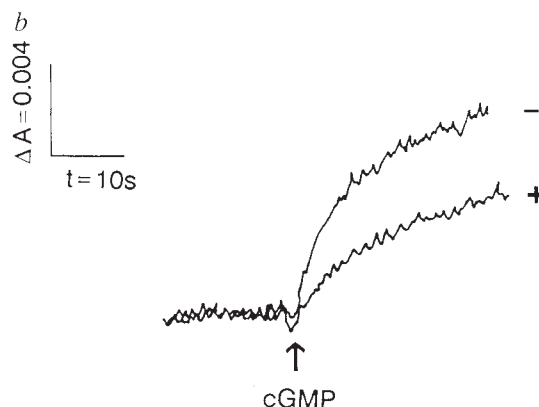
FIG. 1 a, Effect of calmodulin on the activation kinetics of the cGMP-gated channel. The relative initial velocity of channel activation by cGMP was plotted against the concentration of cGMP either in the presence (■) or absence (●) of calmodulin. In the absence of calmodulin, the channel had an apparent K_m of $19 \mu\text{M}$ and a Hill coefficient of 3.8 for cGMP. In the presence of calmodulin, its K_m increased to $33 \mu\text{M}$, but its Hill coefficient remained the same at 3.8. These assays were done in three separate experiments using different ROS membrane vesicle preparations; in all cases an increase in K_m (from $19 \pm 0.4 \mu\text{M}$ to $33 \pm 2 \mu\text{M}$) in the presence of calmodulin without significant change in either the V_{max} or the Hill coefficient ($n = 3.7 \pm 0.1$ and $n = 3.5 \pm 0.6$ in the absence and presence of calmodulin, respectively) was observed. The solid curved lines are calculated from a sigmoidal binding isotherm¹¹ using the indicated K_m and Hill coefficients. b, A typical trace for cGMP-dependent influx of Ca^{2+} into ROS vesicles containing trapped Arsenazo III. A cGMP concentration of $12.5 \mu\text{M}$ was used to initiate the influx of Ca^{2+} either in the presence (+) or absence (–) of calmodulin.

METHODS. ROS membrane vesicles were loaded with Arsenazo III dye as follows: ROS membranes were washed twice in buffer A (10 mM HEPES pH 7.4, 1 mM DTT, 1 mM EDTA and 0.1 mM diisopropylfluorophosphate (DFP)) and once in buffer B (2 mM HEPES, pH 7.4, 0.2 mM DTT and 0.02 mM DFP) to remove soluble proteins. The membrane pellet was then resuspended in

Ca^{2+} -calmodulin was observed to increase the K_m for cGMP from $34 \mu\text{M}$ to $51 \mu\text{M}$ without affecting the Hill coefficient ($n = 2.2$). The higher K_m and lower Hill coefficient obtained with this assay may be due to the hydrophobic nature of this dye and its interaction with the phospholipids and/or channel.

The effect of Ca^{2+} concentration on the calmodulin modulation of channel activity was also measured using the neutral red dye assay. As shown in Fig. 3 for Na^+ influx into ROS membrane vesicles, the calmodulin effect on channel activity was observed over a Ca^{2+} concentration of 50–300 nM. This is comparable to the range of the Ca^{2+} dependence for recoverin activation of guanylate cyclase^{5–7} and to the physiological Ca^{2+} concentration in photoreceptor outer segments^{15–18}. Earlier, Caretta *et al.*¹⁹ observed that Ca^{2+} decreased the binding affinity of a cGMP analogue to ROS membranes. It is likely that this observed Ca^{2+} effect on cGMP binding is related to the effect of Ca^{2+} on calmodulin modulation of the cGMP-gated channel as reported here.

A 240K protein previously shown to be tightly associated with the cGMP-gated channel of ROS membranes¹⁰ has been identified as a major calmodulin-binding protein of ROS membranes by affinity chromatography and western blotting. Two polypeptides of apparent M_r 63K and 240K were isolated as the major constituents by calmodulin affinity chromatography (Fig. 4 (left), lane b). These polypeptides comigrated with the 63K and 240K polypeptides of the cGMP-gated channel complex isolated on an anti-channel immunoaffinity column^{10,20} (Fig. 4 (left), lane c). Western blots labelled with antibodies to the channel and to the 240K protein¹⁰ confirmed that the calmodulin affinity column extracted the 63K/240K cGMP-gated channel complex from ROS membranes (Fig. 4 (middle)). A western blot labelled with ^{125}I -calmodulin in the presence of Ca^{2+} was



buffer B at a concentration of 12 mg protein per ml. The ROS membrane was then bleached by continuous white light and subjected to freezing and thawing. The membranes suspended in 10 mM HEPES, pH 7.4, 100 mM KCl, 1 mM DTT, and 2 mM Arsenazo III were sonicated for 1 min in a beaker of water using a broad-tip sonicator probe (Heat Systems-Ultrasonics) at setting 7. The sonicated ROS membranes were then extruded three times through a Lipid Extruder (Lipex Biomembranes, Vancouver) using two layers of Nuclepore polycarbonate membranes of decreasing pore sizes from 800 nm to 400 nm and finally to 200 nm (ref. 28). The membrane vesicles were then passed through a Sephadex-G50 column ($1.5 \times 28 \text{ cm}$) equilibrated in buffer C (10 mM HEPES, pH 7.4, 1 mM DTT and 100 mM KCl) to remove untrapped dye. The eluted vesicles were further dialysed against buffer C for 3 h. ROS membrane vesicles containing trapped Arsenazo III dye were then added to buffer C in a cuvette to a final volume of 2 ml and a protein concentration of 0.29 mg ml^{-1} either in the absence or presence of $4 \mu\text{g ml}^{-1}$ (235 nM) calmodulin (Sigma). Stock CaCl_2 (4 μl) was then added to give a final concentrations of $100 \mu\text{M}$ CaCl_2 . One minute after the addition of calcium, the calcium influx assay was initiated by the injection of 4 μl of varying concentrations of cGMP. The change in absorbance was monitored at 650–730 nm using a SLM Aminco DW 2000 dual-wavelength spectrophotometer.

used to identify the calmodulin-binding proteins. Although calmodulin labelled several polypeptides of apparent M_r 240K, 140K, 70K, 67K and 35K in both the ROS membrane and calmodulin affinity-purified fraction (Fig. 4 (right), lanes a and b), only the 240K polypeptide was labelled with calmodulin in the immunoaffinity-purified channel complex (Fig. 4 (right) lane c). Calmodulin affinity chromatography has also been used to purify a related 63K/240K channel complex from pig and rat ROS membranes. This suggests that the channel complex and its interaction with calmodulin is conserved, at least for mammalian ROS. These results indicate that calmodulin affinity chromatography can be used to purify the cGMP-gated channel complex from ROS membranes and that the 240K channel-associated polypeptide binds calmodulin in a Ca^{2+} -dependent manner. On the basis of these findings, calmodulin modulation of the channel activity is likely to occur through Ca^{2+} -dependent interaction of calmodulin with the 240K channel-associated protein. The molecular identity of the 240K protein is not yet known. Previous studies have indicated some immunocross-reactivity between the 240K ROS protein and the α -subunit of red blood cell spectrin and brain fodrin^{10,21}. These proteins have been reported to bind to calmodulin²².

Results reported here suggest that changes in cytoplasmic Ca^{2+} concentration during the photoresponse can affect the activity of the cGMP-gated channel. The channel is opened in response to the cooperative binding of cGMP to the 63K channel subunits. The affinity of the channel for cGMP is further modulated by the level of cytoplasmic Ca^{2+} through the binding of Ca^{2+} -calmodulin to the channel complex. This Ca^{2+} -calmodulin

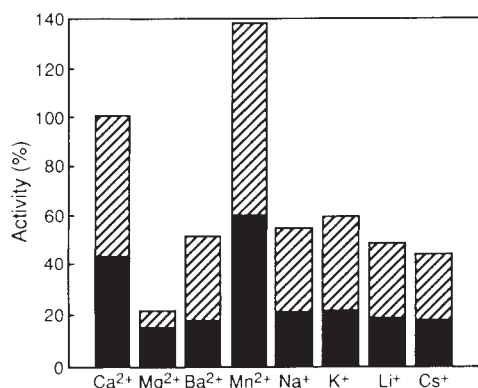


FIG. 2 The initial rate of cGMP-dependent influx of different monovalent (Na^{+} , K^{+} , Li^{+} and Cs^{+}) and divalent (Mg^{2+} , Ba^{2+} and Mn^{2+}) cations into ROS membrane vesicles in the presence (solid) and absence (cross-hatched) of calmodulin. A cGMP concentration of $20 \mu\text{M}$ was used to initiate cation influx. The initial velocity of Ca^{2+} flux in the absence of calmodulin was set at 100%. All ions except Mg^{2+} showed a two- to threefold decrease in initial velocities in the presence of Ca^{2+} -calmodulin.

METHODS. ROS were washed once in buffer A (10 mM HEPES/arginine, pH 7.4, 0.5 mM DTT, 1 mM EDTA) and once in buffer B (10 mM HEPES/arginine, pH 7.4 and 0.5 mM DTT) to remove soluble proteins. The membrane pellet was then resuspended in buffer B at a concentration of 14 mg protein per ml and then subjected to freezing and thawing. The membranes were then mixed with an equal volume of the resuspension buffer (20 mM HEPES/arginine, pH 7.4, 10% sucrose, 2.5% Ficoll, 0.5 mM DTT) and extruded four times through the lipid extruder using one layer of Nuclepore polycarbonate membrane of decreasing pore size from 800 nm to 400 nm. The assay was modified from ref. 13. The membrane vesicles were added to the assay buffer containing 20 mM HEPES/arginine, pH 7.4, 5% sucrose, 0.5 mM DTT, 30 μM neutral red, 1 μM trifluoromethoxycarbonylcyanide phenylhydrazide and with or without 400 nM calmodulin to make up a total volume of 2 ml and a concentration of 0.5 mg protein per ml. The assay was initiated by the addition of 50 μl of 0.5 M of various ions (12.2 mM final concentration) containing 0.4 mM CaCl_2 (10 μM final concentration), followed by the addition of 4 μl 10 mM cGMP (20 μM final concentration) 1 min later. The change in absorbance was monitored at 540–650 nm using a SLM Aminco DW 2000 dual-wavelength spectrophotometer.

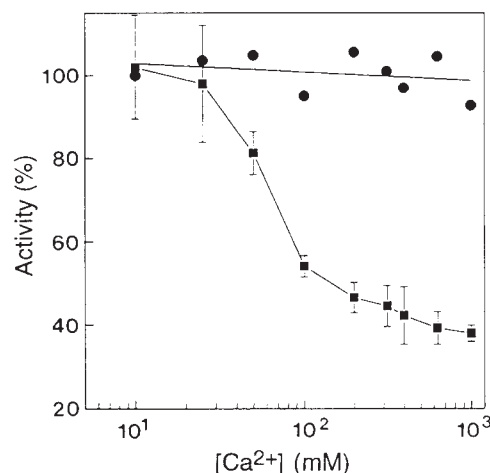


FIG. 3 Calcium dependence of calmodulin effect on the cGMP-gated channel activity. The initial velocity of Na^{+} influx in the presence of 10 nM Ca^{2+} but in the absence of calmodulin was designated as 100% activity. The assay was done either in the presence (■) or absence (●) of calmodulin and with 20 μM of cGMP. Results of three sets of experiments in the presence of calmodulin are shown with standard deviations marked by error bars.

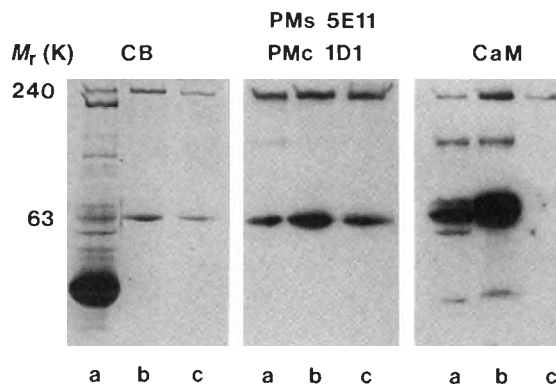
METHODS. Assay procedures are as described for Fig. 2. Assay buffer containing different concentrations of CaCl_2 was prepared according to the program of Fabiato²⁹. The following ratios of Ca^{2+} to 1 mM EGTA were used (estimated free Ca^{2+} in parentheses): 0.134 (10 nM), 0.279 (25 nM), 0.436 (50 nM), 0.607 (100 nM), 0.755 (200 nM), 0.830 (316 nM), 0.860 (398 nM), 0.908 (631 nM) and 0.940 (1,000 nM).

effect on the channel is most pronounced at low cGMP concentrations as found under physiological conditions²³. Regulation of the channel by cGMP and Ca^{2+} -calmodulin is in some ways analogous to the regulation of haemoglobin by oxygen and pH (Bohr effect), in which small changes in pH alter the affinity of haemoglobin for oxygen.

The Ca^{2+} -calmodulin effect on the cGMP-gated channel can be incorporated into the current model for photoexcitation and photorecovery^{1-4,18}. In the dark, the free cGMP level in ROS, estimated to be about 4–10 μM , maintains a small but significant number of channels in their open state²³, allowing for the influx of Na^{+} and Ca^{2+} . Under these conditions, the cytoplasmic Ca^{2+} is maintained at a relatively high concentration ($\sim 0.3 \mu\text{M}$) by balancing the influx of Ca^{2+} through the channel with the efflux of Ca^{2+} through the $\text{Na}^{+}/\text{Ca}^{2+}$ -K⁺ exchanger. Under these conditions, guanylate cyclase is maintained at its basal level of activity and the cGMP-gated channel is in its low affinity (high K_m) state for cGMP through Ca^{2+} -calmodulin binding to the 240K channel-associated protein. Photoexcitation leads to the activation of phosphodiesterase and a decrease in cGMP levels. The decrease in total cGMP has been reported to be only 10–15% (ref. 24), but the decrease in free cGMP has not yet been measured. Because the channel is in its low-affinity state, it may be positioned on the cGMP dose-response curve to be sensitive to a small decrease in free cGMP concentration, thus facilitating its closure. A decrease in cytoplasmic Ca^{2+} levels (estimated to be below 100 nM) will result because of the decreased influx of Ca^{2+} through the channel. This decrease in Ca^{2+} will activate the guanylate cyclase⁵, possibly through recoverin^{6,7} and limit phosphodiesterase activation²⁵. Calmodulin will also dissociate from the channel complex and cause the channel to switch to its high-affinity state for cGMP. The channel will now reopen at a lower cGMP level, thus aiding the recovery of the ROS to its dark level as cGMP synthesis proceeds. The opening of the channel will in turn restore the Ca^{2+} level to its dark level, leading to the inactivation of guanylate cyclase and the return of the channel to its low-affinity state. Thus, reversible binding

FIG. 4 Purification of the cGMP-gated channel complex by calmodulin and immunoaffinity chromatography and identification of the 240K channel-associated protein as a major calmodulin-binding polypeptide of ROS membranes. Lanes a, ROS membranes after the extraction of soluble proteins; lanes b, EDTA eluant from a calmodulin affinity column; and lanes c, N-terminal peptide eluant from an N-terminal-specific anti-channel monoclonal antibody (PMc 6E7) column. Left, SDS gel stained with Coomassie blue (CB); middle, western blot labelled with an anti-63K channel monoclonal antibody (PMc 1D1) and an anti-240K monoclonal antibody (PMs 5E11). Right, western blot labelled with 125 I-calmodulin (CaM). Calmodulin was found to bind to the 240K polypeptide, but not to the 63K channel subunit in the immunoaffinity-purified channel complex. No labelling was observed in the presence of EGTA. The intense labelling of the 67K and 70K proteins in the calmodulin column eluant as compared to the 240K protein is probably due to the low transfer efficiency of the 240K protein.

METHODS. ROS membranes were prepared from freshly dissected bovine retinae as previously described¹⁰ in the presence of proteolytic inhibitors: 0.1 mM DFP, 5 μ g ml⁻¹ aprotinin, 2 μ g ml⁻¹ leupeptin, and 2 μ g ml⁻¹ pepstatin. The ROS membranes were washed three times in 10 mM HEPES, pH 7.4 buffer containing 1 mM EDTA and solubilized in solubilization buffer (10 mM HEPES, pH 7.4, 150 mM KCl, 10 mM CaCl₂, 1 mM DTT, 18 mM 3-((3-cholamidopropyl)-dimethylammonio)-1-propanesulphonate (CHAPS), and 0.2% asolectin) at a concentration of 1.5–2 mg protein per ml. Solubilized protein (10 mg) was then loaded onto each ml of calmodulin-Sepharose column (Sigma) equilibrated in the solubilization buffer. The column was then washed with 15 column volumes of washing buffer (10 mM HEPES, pH 7.4, 1 mM DTT, 1 mM CaCl₂, 15 mM CHAPS and 0.18% asolectin). Calmodulin-binding proteins were then eluted off the column with the washing buffer in which 1 mM CaCl₂ was substituted by 2 mM EDTA. For the antibody affinity chromatography, PMc 6E7 monoclonal antibody was purified from



ascites fluid and coupled to CNBr-activated Sepharose 2B beads¹⁰. The purification of the cGMP-gated channel complex was essentially the same as the calmodulin affinity chromatography except DTT was omitted and 0.9 mg ml⁻¹ of synthetic peptide corresponding to the N-terminal of the 63K ROS channel (Ser-Asn-Lys-Glu-Gln-Glu-Pro-Lys-Glu-Lys-Lys-Lys) was used to elute the channel complex²⁰. SDS gel electrophoresis and western blotting were done as described¹⁰. Calmodulin was iodinated by Bolton-Hunter reagent (NEN) and the labelling of the blot by calmodulin was done according to the method of Flanagan *et al.*³⁰. For immunoblotting analysis, the blot was first labelled with PMc 6E7 and PMs 5E11 and then relabelled with 125 I-labelled goat anti-mouse immunoglobulin and visualized by autoradiography.

of calmodulin to the channel complex in response to changes in the level of Ca²⁺ has the effect of increasing the sensitivity of the channel to small changes in cGMP which may occur during the photoresponse.

Light adaptation characterized by a decrease in rod sensitivity in the presence of background illumination has been shown to be mediated by Ca²⁺ (refs 26, 27). Reduced Ca²⁺ induced under conditions of light adaptation will result not only in an increase in cGMP⁵, but also in a higher affinity of the channel for cGMP. On this basis, Ca²⁺ may regulate photorecovery and light adaptation not only by regulating guanylate cyclase activity^{6,7} and phosphodiesterase activation²⁵, but also by modulating the affinity of the channel for cGMP. □

T-cell adhesion induced by proteoglycan-immobilized cytokine MIP-1 β

Yoshiya Tanaka*, David H. Adams*, Stefan Hubscher†, Hiroyuki Hirano*, Ulrich Siebenlist‡ & Stephen Shaw*

* The Experimental Immunology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

† Department of Pathology, University of Birmingham, Birmingham B15 2TT, UK

‡ The Laboratory of Immunoregulation, National Institutes of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

LYMPHOCYTE migration from blood into tissue depends on integrin-mediated adhesion to endothelium^{1–4}. Adhesion requires not only integrin ligands on the endothelium, but also activation signals because T-cell integrins cannot bind well until they are activated. The physiological 'triggers' for T-cell adhesion are unknown, but cytokines may be good candidates as they are released during inflammation and trigger adhesion in neutrophils and monocytes^{2,5,6}. We have identified a cytokine, macrophage inflammatory protein-1 β (MIP-1 β), that induces both chemotaxis and adhesion of T cells; MIP-1 β is most effective at augmenting adhesion of CD8⁺ T cells to the vascular cell adhesion molecule VCAM-1. We reasoned that, as cytokines *in vivo* will be rapidly washed away, MIP-1 β might be bound to endothelial surfaces and so induce adhesion in its immobilized form. Here we show that: (1) MIP-1 β is present on lymph node endothelium; (2) immobilized MIP-1 β induces binding of T cells to VCAM-1 *in vitro*. MIP-1 β was immobilized by binding to proteoglycan: a conjugate of heparin with bovine serum albumin and cellular proteoglycan CD44 were both effective. We propose that MIP-1 β and other cytokines with glycosaminoglycan-binding sites will bind to and be presented by endothelial proteoglycans to trigger adhesion selectively not only of lymphocyte subsets, but also of other cell types.

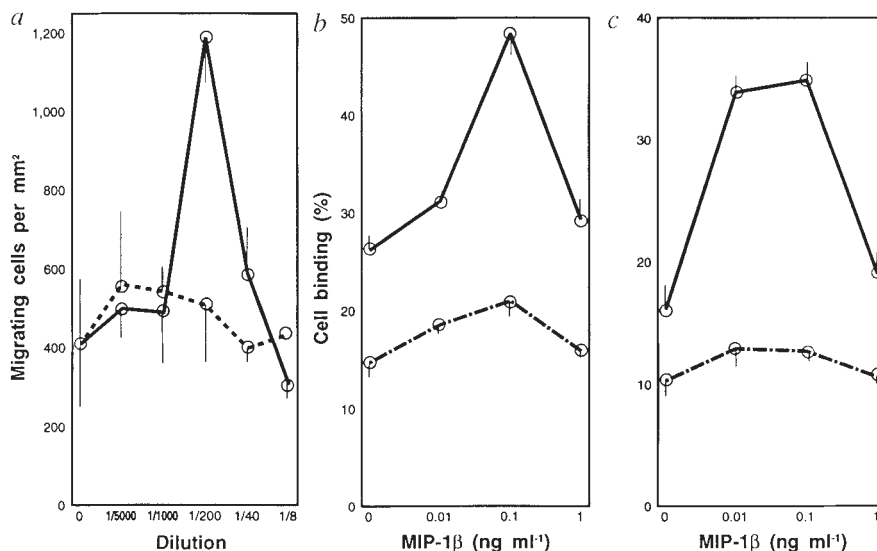
Received 20 July; accepted 23 October 1992.

1. Stryer, L. A. *Rev. Neurosci.* **9**, 87–119 (1986).
2. Chabre, M. & Deterre, P. *Eur. J. Biochem.* **179**, 255–266 (1989).
3. Pugh, E. N. Jr. & Lamb, T. D. *Vision Res.* **30**, 1923–1948 (1990).
4. Yau, K.-W. & Baylor, D. A. *Rev. Neurosci.* **12**, 289–328 (1989).
5. Koch, K.-W. & Stryer, L. *Nature* **334**, 64–66 (1988).
6. Ditzhoor, A. M. *et al. Science* **251**, 915–918 (1991).
7. Lambrecht, H.-G. & Koch, K.-W. *EMBO J.* **10**, 793–798 (1991).
8. Kohnen, R. E., Chafouleas, J. G., Eadie, D. M., Means, A. R. & McConnell, D. G. *J. Biol. Chem.* **256**, 12517–12522 (1981).
9. Nagao, S., Yamazaki, A. & Bitensky, M. W. *Biochemistry* **26**, 1659–1665 (1987).
10. Molday, L. L., Cook, N. J., Kaupp, U. B. & Molday, R. S. *J. Biol. Chem.* **265**, 18690–18695 (1990).
11. Koch, K.-W. & Kaupp, U. B. *J. Biol. Chem.* **260**, 6788–6800 (1985).
12. Malencik, D. A. & Anderson, S. R. *Biochem. biophys. Res. Commun.* **114**, 50–56 (1983).
13. Schnetkamp, P. P. M. *J. gen. Physiol.* **96**, 517–534 (1990).
14. Ogawa, Y. & Tanokura, M. *J. Biochem.* **95**, 19–28 (1984).
15. McNaughton, P. A., Cervetto, L. & Nunn, B. J. *Nature* **322**, 261–263 (1986).
16. Ratto, G. M., Payne, R., Owen, W. G. & Tsien, R. Y. *J. Neurosci.* **8**, 3240–3246 (1988).
17. Korenbrok, J. I. & Miller, D. L. *Vis. Res.* **29**, 939–948 (1989).
18. Kaupp, U. B. & Koch, K.-W. *A. Rev. Physiol.* **54**, 153–175 (1992).
19. Caretta, A., Cavagioni, A., Gimaldi, R. & Sorbi, R. T. *Eur. J. Biochem.* **177**, 139–146 (1988).
20. Molday, R. S. *et al. J. Biol. Chem.* **266**, 21917–21922 (1991).
21. Wong, S. & Molday, R. S. *Biochemistry* **25**, 6294–6300 (1986).
22. Glenney, J. R. Jr, Glenney, P. & Weber, K. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4002–4005.
23. Nakatani, K. & Yau, K.-W. *J. Physiol.* **395**, 695–729 (1988).
24. Cote, R. H., Nicol, G. D., Burke, S. A. & Bownds, M. D. *J. Biol. Chem.* **261**, 12965–12975 (1986).
25. Kawamura, S. & Murakami, M. *Nature* **349**, 420–423 (1991).
26. Torre, V., Matthews, H. R. & Lamb, T. D. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7109–7113 (1986).
27. Nakatani, K. & Yau, K.-W. *Nature* **334**, 69–71 (1988).
28. Hope, M. J., Bally, M. B., Webb, G. & Cullis, P. R. *Biochim. biophys. acta* **812**, 55–65 (1985).
29. Flanigan, A. *Meth. Enzym.* **157**, 378–417 (1988).
30. Flanagan, S. D. & Yost, B. *Analyt. Biochem.* **140**, 510–519 (1984).

ACKNOWLEDGEMENTS. We thank P. Schnetkamp for discussions related to the neutral red assay, S.-C. Hsu and W. Rodriguez for technical advice and A. F. X. Goldberg for discussions. This work was supported by a grant from the NIH. Y.-T.H. is supported by an MRC studentship.

FIG. 1 MIP-1 β induces T-cell chemotaxis and adhesion of T cells to VCAM-1 and fibronectin. *a*, Chemotactic response of resting T cells to increasing concentrations of MIP-1 β (solid line) or MIP-1 α (broken line). *b* and *c*, Adhesion of CD4 T cells (long-dash line) and CD8 T cells (solid line) to VCAM-1 (*b*) and FN (*c*) in the continuous presence of increasing concentrations of recombinant human MIP-1 β .

METHODS. T cells, CD4⁺ T cells and CD8⁺ T cells were prepared by exhaustive negative selection from peripheral blood mononuclear cells of normal donors using immunomagnetic beads and multi-antibody cocktail as described³⁴. Analysis of expression of markers (CD4, CD8) on the relevant subsets always showed >96% purity. We use the term MIP-1 β , originally applied to the mouse gene product, for its human counterpart used here. *a*, Cell chemotaxis was assessed in 48-well microchemotaxis chambers as described³⁵ using 5- μ m polycarbonate chemotaxis membranes (Nuclepore) coated by floating overnight on type IV collagen (5 μ g ml⁻¹) (Sigma)³⁶. Cells migrating to the underside of the membrane were counted with an Optomax System IV image analyser. MIP-1 β was diluted from cell supernatants from conditioned media collected from SF9 insect cell lines that had been infected with baculovirus in which pAT744 MIP-1 β cDNA was expressed²⁴; MIP-1 α was prepared similarly, except that pAT 464 was substituted²⁴. All samples were assayed in triplicate and the mean value used; lines represent standard errors. *b* and *c*, Adhesion was assayed as described³⁷. Purified VCAM-1 (35 ng per well; from W. Newman)¹⁰ or fibronectin (1 μ g per well; New York Blood Center) was applied to 96-well microtitre



plates (Costar). ⁵¹Cr-labelled T cells and recombinant human MIP-1 β (from M. Tsang) were added to the plates, settled at 4 °C for 60 min and rapidly warmed to 37 °C for 15 min, after which non-adherent cells were washed off. γ -Emissions of the lysates of adherent cells were counted. Background binding of T cells to BSA or collagen was less than 3%. Data are expressed as a mean percentage of the binding of T cells from a representative experiment; lines represent standard errors.

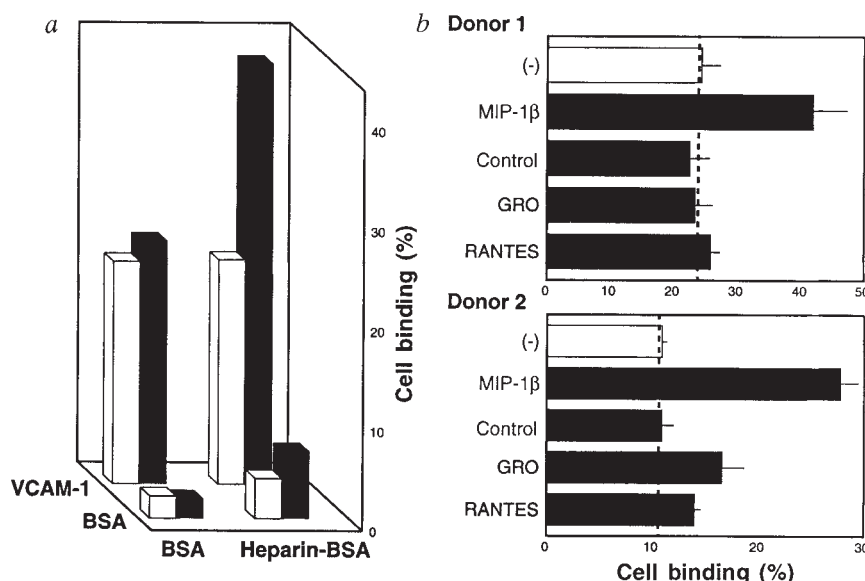
From a survey of cytokines, we found that one, MIP-1 β , was unique in mediating both chemotaxis and adhesion-induction for T cells (Fig. 1). MIP-1 β is a member of the 'inter-crine'/'chemokine' family, which includes chemotactic cytokines such as the interleukin IL-8, MCP and RANTES^{7,8}. We observed a chemotactic response to MIP-1 β over a narrow range of cytokine concentrations (Fig. 1*a*; and data not shown); MIP-1 β has been reported to induce myeloid cell chemotaxis at higher concentrations⁹ and apparently T-cell chemotaxis^{2,8}. MIP-1 β also induces augmented adhesion of CD8 T cells to VCAM-1 over a similar concentration range (Fig. 1*b*); this

augmentation is superimposed on the basal binding to VCAM-1 with resting T cells¹⁰. MIP- β also increases binding of CD8 cells to the extracellular matrix ligand fibronectin (Fig. 1*c*), although, unlike adhesion to VCAM-1, this response is not observed in all donors (data not shown). MIP-1 β augmentation of adhesion occurs predominantly with CD8⁺ rather than CD4⁺ cells (Fig. 1*b, c*). The induction of adhesion of particular T-cell subsets by specific cytokines would make the process of lymphocyte recruitment more flexible and selective^{2,3,11}.

We and others^{12,13} have proposed that cytokines should be most effective when immobilized on the luminal surface of the

FIG. 2 Immobilized MIP-1 β retained by heparin-BSA induces CD8⁺ T-cell adhesion to co-immobilized VCAM-1 but other cytokines do not. *a*, MIP-1 β (filled bars) or control medium (open bars) was added to wells which had been coated with one of four possible combinations of VCAM-1 or BSA and heparin-BSA or BSA. The cytokine was washed out before CD8⁺ T cells were added and binding was evaluated as for Fig. 1. Increased binding in the presence of MIP-1 β was seen in each of 21 donors in 13 experiments, with an overall statistical significance of $P < 0.001$ (Student's *t*-test). *b*, MIP-1 β induces adhesion in the presence of heparin-BSA but GRO and RANTES do not. Open bar and dotted line define the level of binding of T cells to heparin-BSA/VCAM-1 in the absence of cytokines.

METHODS. Purified VCAM-1 (35 ng per well), BSA-heparin conjugates³⁸ (20 ng per well; from D. Roberts) and/or control BSA in PBS were applied to plastic plates at 4 °C overnight and the plastic binding sites subsequently blocked with BSA. MIP-1 β (1:100), control supernatant from SF9 cells infected with wild-type baculovirus (1:100), GRO/MGSA (10 ng ml⁻¹; from J. J. Oppenheim), or RANTES (10 ng ml⁻¹; from J. J. Oppenheim) were added to the wells and allowed to bind to plastic-immobilized heparin-BSA at 37 °C for 60 min. Free cytokine was then removed by washing the plates 5 times in PBS; ⁵¹Cr-labelled CD8⁺ T cells were added to the plates and the number of cells



binding was assayed as described in Fig. 1 legend. Data are expressed as mean percentage of T cells binding from three replicate wells; lines represent standard errors.

endothelium, otherwise blood flow would prevent accumulation of effective local concentrations. Chemokines like MIP-1 β have glycosaminoglycan (GAG)-binding sites and bind heparin^{2,7,8,14}. There are many biological systems in which the interaction of cytokines with proteoglycans in the extracellular matrix^{15,16} and on cell surfaces^{15,17,18} is functionally important. We propose a novel variation, namely that proteoglycans on an endothelial cell can bind and present cytokines to passing leukocytes.

We have therefore created a biochemically defined model system to explore this hypothesis. (1) Heparin-bovine serum albumin (BSA) conjugate (a model proteoglycan) and VCAM-1 (an endothelial ligand for VLA-4) were co-immobilized on plastic; (2) MIP-1 β was added and then washed out to assess only the effect of immobilized cytokine; and (3) we assayed binding of resting CD8⁺ T cells to all eight combinations of the three components: integrin ligand, proteoglycan and cytokine (Fig. 2a). A marked MIP-1 β -specific augmentation of binding was observed only when both the proteoglycan and the integrin ligand were co-immobilized. Other cytokines bearing GAG-binding sites such as RANTES and GRO showed minimal activity (Fig. 2b; and data not shown). Thus MIP-1 β can be retained by immobilized proteoglycan and induce T-cell adhesion to integrin ligands on that same surface.

To explore whether cell surface proteoglycan could also present MIP-1 β to T cells, we substituted CD44 for heparin-BSA conjugate in our model. CD44, a proteoglycan^{19,20} expressed on many cells, including endothelium²¹, has been implicated in T-cell adhesion to endothelial cells²². VCAM-1 and purified CD44 glycoprotein were co-immobilized, MIP-1 β added, incubated, free cytokine washed out, and the binding of T cells assayed (Fig. 3a). The results parallel those with synthetic proteoglycan. MIP-1 β -specific augmentation of binding was consistently observed only when CD44 and the integrin ligand VCAM-1 were co-immobilized. MIP-1 β -specific augmentation of adhesion to VCAM-1 was not observed with a control glycoprotein (CD45; data not shown). Blocking studies with mono-

clonal antibodies confirmed that adhesion was mediated by T-cell VLA-4 binding to VCAM-1 (Fig. 3b). The presence of heparin and heparan sulphate during the phase of MIP-1 β immobilization substantially inhibited subsequent adhesion induction (Fig. 3c), suggesting an involvement for GAG-binding.

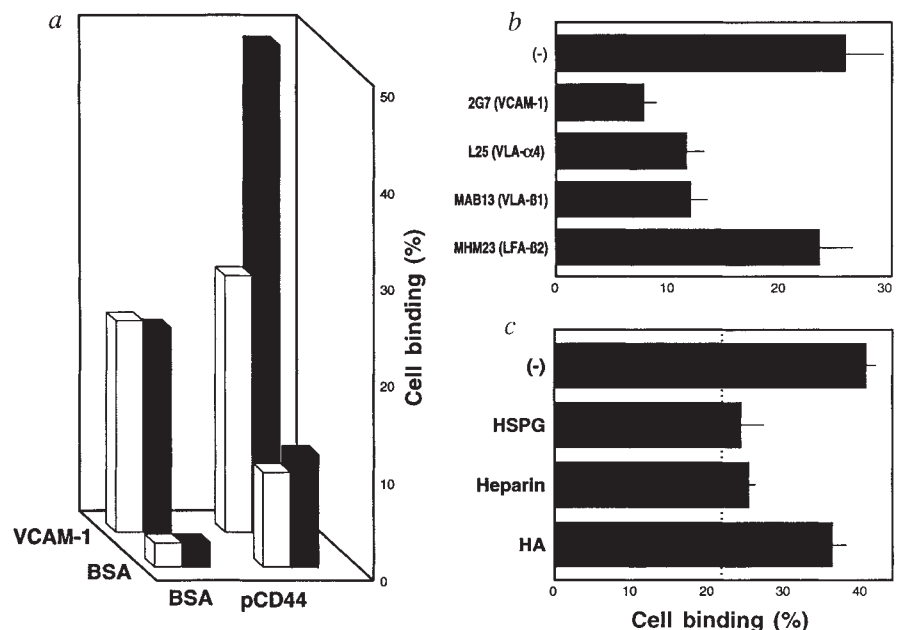
MIP-1 β is produced in large amounts by monocytes, fibroblasts and lymphocytes early after activation *in vitro*^{23,24}; *in vivo*, it might bind to proteoglycan on the endothelial surface at inflammatory sites and be presented as an adhesion-inducing stimulus to passing leukocytes. Our hypothesis is supported by the detection of MIP-1 β on endothelium in lymphoid tissue (Fig. 4) as well as on endothelium at sites of inflammation such as tonsil (data not shown). Interleukin-8 induces granulocyte adhesion^{25,26} and can be localized at the surface of endothelium¹², findings that taken together with our MIP-1 β results indicate that endothelial presentation of pro-adhesive cytokines may be a general mechanism¹³.

Endothelial cells express many proteoglycans, whose structure and density depend on the type of endothelium and its activation status^{14,27,28}. Our data indicate that cytokines are among the many proteins that bind to endothelial surfaces in a GAG-dependent fashion¹⁴. CD44, which we have used as a model proteoglycan, is a plausible candidate for *in vivo* presentation of pro-adhesive cytokines; such a role for endothelial CD44 would be additional to the apparent role of T cell CD44 in T-cell binding to endothelium²⁹. CD44 and other endothelial proteoglycans show structural diversity in both their protein core and their carbohydrate side chains^{14,30}; this could introduce additional specificity to cytokine presentation during adhesion cascades^{2,13}.

A model adhesion cascade has been proposed to explain adhesion of many leukocyte subtypes to endothelium (reviewed in refs 2, 3, 31, 32). Circulating cells are first loosely 'tethered' to endothelium by selectin-mediated adhesion. This would allow presentation of proteoglycan-bound MIP-1 β by endothelial cells

FIG. 3 MIP-1 β retained by purified CD44 protein induces CD8⁺ T-cell adhesion to co-immobilized VCAM-1 in a fashion that is dependent on VLA-4/VCAM-1 and GAG. **a**, Experimental design is identical to that for Fig. 2a but uses purified CD44 in place of heparin-BSA. MIP-1 β (filled bars) or control (open bars) was added to the wells and washed away before assaying adhesion. Increased binding in the presence of MIP-1 β has been seen in each of 11 donors in 7 experiments, with an overall statistical significance of $P < 0.001$ (Student's *t*-test). **b**, Inhibition of CD44-immobilized MIP-1 β -induced T-cell adhesion to VCAM-1 by monoclonal antibodies against $\beta 1$ integrins and VCAM-1. **c**, Effect of free GAGs on T-cell adhesion to VCAM-1 induced by CD44-immobilized recombinant MIP-1 β . GAGs were applied to plates before addition of MIP-1 β and washed away when MIP-1 β was removed. Dotted line indicates the level of binding of T cells to CD44/VCAM-1 in the absence of MIP-1 β .

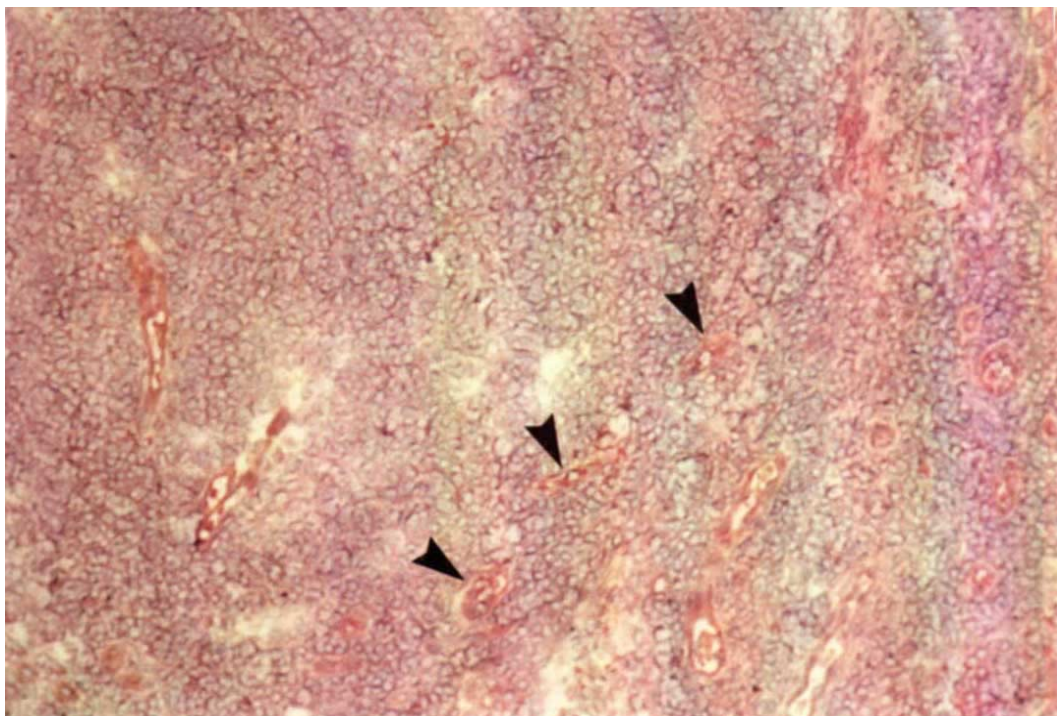
METHODS. CD44 glycoprotein was purified from a monocyte-rich preparation of peripheral blood mononuclear leukocytes by immunoaffinity chromatography using monoclonal antibody raised against CD44, NIH44-1 (ref. 39), and applied at ~ 1 ng per well (estimated by silver staining). The preparation consists of NIH44-1-immunoreactive components with M_r s estimated as dominant ~ 80 – 90 K and minor ~ 200 – 220 K by western blotting after transfer from a reducing SDS-polyacrylamide gel; VCAM-1 and/or control BSA were co-immobilized as for Fig. 2. MIP-1 β (1:100) or control conditioned medium was added to plastic co-immobilized protein for 30 min, free cytokine was then washed out and adhesion was assayed as for Fig. 2. In **b**, the following monoclonal antibodies were added at saturating concentrations of $10 \mu\text{g ml}^{-1}$: anti-VCAM-1 mAb 2G7 (from



W. Newman); VLA-4 (CD49d) mAb L25 (Becton-Dickinson); $\beta 1$ (CD29) mAb MAB13 (from K. Yamada) and CD18 mAb MHM23 (from J. E. Hildreth). In **c**, MIP-1 β was used at 10 ng ml^{-1} final concentration. The following free GAGs were added to wells before addition of MIP-1 β : heparan sulphate proteoglycan (HSPG) ($30 \mu\text{g ml}^{-1}$; Collaborative Research), heparin sodium (15 U ml^{-1} ; LyphoMed) and hyaluronic acid (HA) ($100 \mu\text{g ml}^{-1}$; Sigma). GAGs and MIP-1 β were removed at the same time by extensive washing.

FIG. 4 Presence of MIP-1 β on lymph node endothelium detected by immunohistochemistry. Many vessels show positive endothelial staining, including some with the morphology of high endothelial venules (arrowed).

METHODS. Frozen sections were stained from the paracortical region of a human inguinal lymph node with features of nonspecific reactive hyperplasia. A three-step immunalkaline-phosphatase method⁴⁰ was used with a rabbit anti-MIP-1 β polyclonal antibody diluted 1/5,000 in Tris-buffered saline. The anti-serum was raised against a synthetic peptide from the C-terminal region of MIP-1 β ; of the many molecules assayed for reactivity, only MIP-1 β was reactive.



to its receptor on T cells. Cytokines, when immobilized, become analogous to previously described triggers that are integral membrane molecules on activated endothelium: the lipid mediator PAF³¹ and E-selectin³³. Such triggers cause functional activation of leukocyte integrins, which then mediate strong adhesion to endothelial ligands: cells can then migrate into tissue under the influence of local chemotactic factors. Extrapolation to include different cytokines and endothelial surface proteoglycans would bring enormous flexibility and selectivity to the process of lymphocyte recruitment¹³. Pharmacological agents could then be selected to modify these cascades. □

36. Larsen, C. G., Anderson, A. O., Appella, E., Oppenheim, J. J. & Matsushima, K. *Science* **243**, 1464–1466 (1989).
37. Tanaka, Y. *et al. J. exp. Med.* **176**, 245–253 (1992).
38. Guo, N. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 3040–3044 (1992).
39. Shimizu, Y., van Seventer, G. A., Siraganian, R., Wahl, L. & Shaw, S. *J. Immun.* **143**, 2457–2463 (1989).
40. Adams, D. H., Hubscher, S. G., Shaw, J., Rothlein, R. & Neuberger, J. M. *Lancet* **2**, 1122–1125 (1989).

ACKNOWLEDGEMENTS. We thank our volunteer blood donors; the NIH Blood Bank; P. Henkart, D. Roberts, T. Schweighoffer and Y. Shimizu for critical review of the manuscript; W. Newman, J. Oppenheim, D. Roberts and M. Tsang for reagents; P. Beverley, D. Kawahara, J. Hildreth, W. Newman, A. von dem Borne, K. Yamada and H. Zola for mAbs; and G. Ginther-Luce for technical assistance. Y.T. and H.H. are Visiting Fellows supported by the Fogarty Exchange Program, and D.H.A. is supported by an Eli Lilly-Medical Research Council travelling fellowship.

Received 6 July 1992; accepted 23 October 1992.

1. Springer, T. A. *Nature* **346**, 425–434 (1990).
2. Butcher, E. C. *Cell* **67**, 1033–1036 (1991).
3. Shimizu, Y., Newman, W., Tanaka, Y. & Shaw, S. *Immun. Today* **13**, 106–112 (1992).
4. Hynes, R. O. *Cell* **69**, 11–25 (1992).
5. Shaw, L. M. & Mercurio, A. M. *J. exp. Med.* **169**, 303–308 (1989).
6. Arnaout, M. A., Wang, E. A., Clark, S. C. & Sieff, C. A. *J. clin. Invest.* **78**, 591–601 (1986).
7. Oppenheim, J. J., Zachariae, C. O. C., Mukaida, N. & Matsushima, K. *A. Rev. Immun.* **9**, 617–648 (1991).
8. Schall, T. J. *Cytokine* **3**, 165–183 (1991).
9. Wolpe, S. D. *et al. J. exp. Med.* **167**, 570–581 (1988).
10. van Seventer, G. A. *et al. J. exp. Med.* **174**, 901–913 (1991).
11. Schall, T. J., Bacon, K., Toy, K. J. & Goeddel, D. V. *Nature* **347**, 669–671 (1990).
12. Rot, A. *Immun. Today* **13**, 291–294 (1992).
13. Tanaka, Y., Adams, D. H. & Shaw, S. *Immun. Today* (in the press).
14. Jackson, R. L., Busch, S. J. & Cardin, A. D. *Physiol. Rev.* **71**, 481–539 (1991).
15. Ruoslahti, E. & Yamaguchi, Y. *Cell* **64**, 867–869 (1991).
16. Nathan, C. & Sporn, M. *J. Cell. Biol.* **113**, 981–986 (1991).
17. Klagsbrun, M. & Baird, A. *Cell* **67**, 229–231 (1991).
18. Gordon, M. Y., Riley, G. P., Watt, S. M. & Greaves, M. F. *Science* **226**, 403–405 (1987).
19. Haynes, B. F., Telen, M. J., Hale, L. P. & Denning, S. M. *Immun. Today* **10**, 423–428 (1989).
20. Brown, T. A., Bouchard, T., St John, T., Wayner, E. & Carter, W. G. *J. Cell Biol.* **113**, 207–211 (1991).
21. Pals, S. T. *et al. J. Immun.* **143**, 851–857 (1989).
22. Berg, E. L. *et al. Immun. Rev.* **108**, 5–18 (1989).
23. Sherry, B. *et al. J. exp. Med.* **168**, 2251–2259 (1988).
24. Zipfel, P. F., Balke, J., Irving, S. G., Kelly, K. & Siebenlist, U. *J. Immun.* **142**, 1582–1590 (1989).
25. Huber, A. R., Kunkel, S. L., Todd, R. F. III & Weiss, S. J. *Science* **254**, 99–102 (1991).
26. Kuijpers, T. W., Hakker, B. C., Hart, M. H. & Roos, D. *J. Cell. Biol.* **117**, 565–572 (1992).
27. Wight, T. N. *Arteriosclerosis* **9**, 1–20 (1989).
28. Kojima, T., Shworak, N. W. & Rosenberg, R. D. *J. Biol. Chem.* **267**, 4870–4877 (1992).
29. Aruffo, A., Stamenkovic, I., Melnick, M., Underhill, C. B. & Seed, B. *Cell* **61**, 1303–1313 (1990).
30. Stamenkovic, I., Aruffo, A., Amiot, M. & Seed, B. *EMBO J.* **10**, 343–348 (1991).
31. Zimmerman, G. A., Prescott, S. M. & McIntyre, T. M. *Immun. Today* **13**, 93–99 (1992).
32. Pardi, R., Inverardi, L. & Bender, J. R. *Immun. Today* **13**, 224–231 (1992).
33. Lo, S. K. *et al. J. exp. Med.* **173**, 1493–1500 (1991).
34. Horgan, K. J. & Shaw, S. in *Current Protocols in Immunology* (eds Coligan, J. E., Kruisbeek, A. M., Margulies, D. H., Shevach, E. M. & Strober, W.) 7.4.1–7.4.5 (Wiley Interscience, New York, 1991).
35. Adams, D. H. *et al. J. Immun.* **147**, 609–612 (1991).

Adhesion of epidermal Langerhans cells to keratinocytes mediated by E-cadherin

Aimin Tang*, Masayuki Amagai*,
Lawrence G. Granger†, John R. Stanley*
& Mark C. Udey*‡

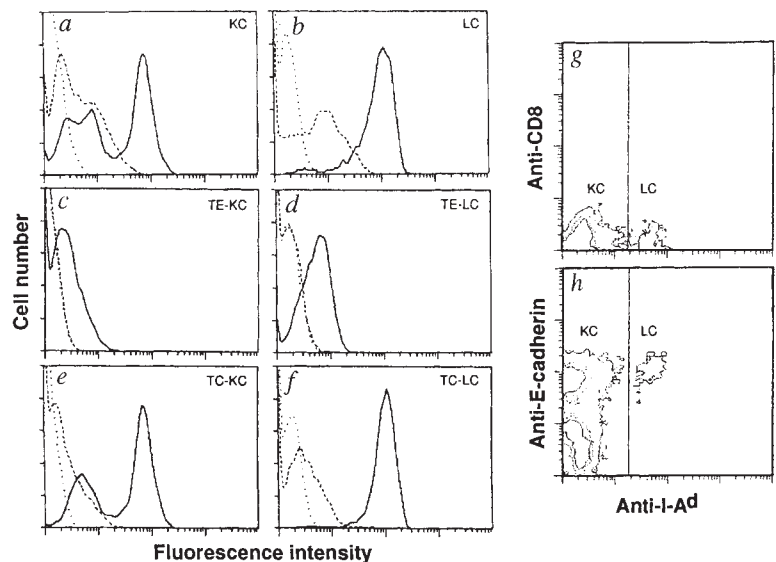
*Dermatology Branch and †Experimental Immunology Branch,
National Cancer Institute, Bethesda, Maryland 20892, USA

LANGERHANS cells (LC) are the principal accessory cells present in epidermis¹. Because LC have limited capacity for self-renewal², epidermis is continually repopulated by as-yet uncharacterized bone marrow-derived LC progenitors^{3,4}. In addition, although LC persist in epidermis for extended periods, LC are induced to migrate from skin to regional lymph nodes after antigen exposure⁵. To begin to elucidate mechanisms involved in LC trafficking, we characterized LC–keratinocyte (KC) interactions. Here we report that fresh murine LC express cadherins, and that LC adhere to KC *in vitro* through E-cadherin. Cultured LC (which may bear a phenotypic and functional relationship to LC that have migrated to lymph nodes^{6,7}) express lower levels of E-cadherin and exhibit

‡ To whom correspondence should be addressed.

FIG. 1 Binding of anti-cadherin monoclonal antibodies to cadherins expressed by epidermal Langerhans cells and keratinocytes. Murine epidermal cells (EC) were treated with trypsin in EDTA (TE) or trypsin in calcium (TC), stained for E- or P-cadherins and I-A antigens, and analysed by two-colour flow cytometry. *a, c, e*, Cadherin-specific (phycoerythrin) fluorescence on I-A⁺ KC; *b, d, f*, phycoerythrin fluorescence on I-A⁺ LC. *a, b*, EC stained after a single exposure to trypsin in calcium during epidermal separation. *c, d, e, f*, TE cells and TC cells, respectively. Profiles of cells stained with rat antibody 53-6.72 (anti-CD8, ATCC, dotted line), ECCD-2 (anti-mouse E-cadherin¹⁴; continuous line) or PCD-1 (anti-mouse P-cadherin⁸; dashed line) are shown. The gating used to differentiate KC stained with anti-CD8 or anti-E-cadherin antibody (as depicted in *a* and *b*) is illustrated in *g* and *h*.

METHODS. Single-cell suspensions of BALB/c EC were prepared as previously described²⁰, except that trypsin digestion was done in Hank's balanced salt solution (HBSS) containing 1 mM calcium (Biofluids), 10 mM HEPES (GIBCO BRL). After dissociation, EC were incubated in 0.01% trypsin in calcium- and magnesium-free HBSS (Biofluids) containing 1 mM EDTA (TE) or 1 mM calcium (TC) for 30 min at 37 °C (ref. 15) and washed with HBSS, 1 mM calcium, 0.005% DNase (Sigma), 20% fetal bovine serum (FBS, Biofluids). Cells were stained for epitopes reactive with rat (6.72, ECCD-2 or PCD-1) and mouse (MKD6, anti-mouse I-A^d; Becton Dickinson) antibody as previously described²¹ and analysed with a FACScan analytical flow cytometer using Consort 30



software (BDIS). LC were >90% viable based on their ability to exclude propidium iodide.

decreased affinity for KC. These results suggest that expression of E-cadherin by LC promotes persistence of these cells in epidermis, and that cadherins may play important and unanticipated roles in interactions between leukocytes and epithelia.

Because cadherins are expressed in uninfamed epidermis⁸ and are prominently involved in interactions between KC⁹⁻¹³, we reasoned that cadherins might also mediate adhesion of LC to KC. We examined murine epidermal cells for surface expression of cadherins with anti-E (ECCD-2; ref. 14) and anti-P-cadherin (PCD-1; ref. 8) monoclonal antibodies by flow cytometry. Staining with ECCD-2 revealed several subpopulations of KC (Fig. 1). About half of the KC expressed high levels

of E-cadherin. The remaining KC expressed lower levels. LC expressed levels of E-cadherin as high as those expressed by the most intensely stained KC. Both KC and LC stained less intensely with PCD-1, but most LC also expressed P-cadherin. We also found that cadherin epitopes on LC (and KC) were degraded by trypsin in 1 mM EDTA (TE cells) and protected from trypsin by 1 mM calcium (TC cells), as previously reported for E- and P-cadherins¹⁵ (Fig. 1). Partially purified splenic dendritic cells and preparations enriched in lymph node dendritic cells did not stain with ECCD-2 or PCD-1 (data not shown).

To demonstrate that the E-cadherin expressed by LC reflected

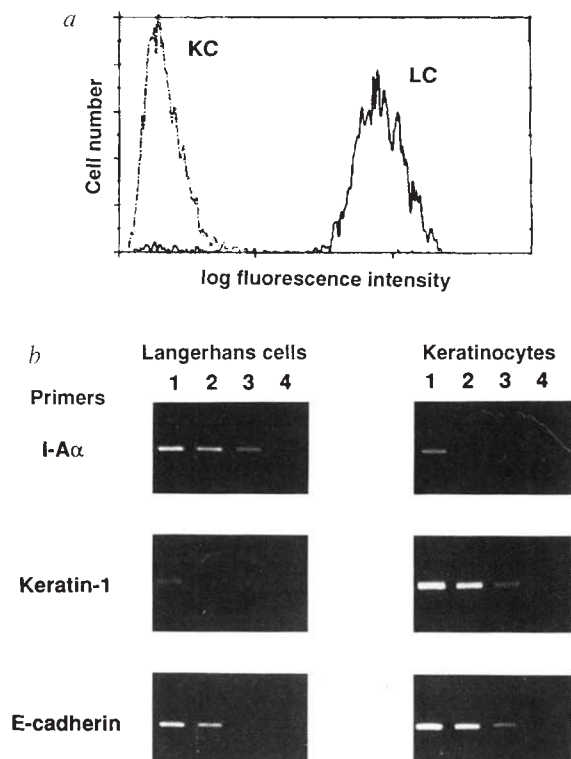


FIG. 2 Characterization of Langerhans cell-derived E-cadherin mRNA. LC were purified to >98% I-A⁺ cells by preparative flow cytometry (*a*, continuous line). KC (<1% I-A⁺) were obtained simultaneously (*a*, interrupted line). Total RNA was prepared and mRNAs were reverse-transcribed and amplified by PCR, resolved in agarose gels and detected after ethidium bromide staining (*b*). Lanes 1-4 depict PCR products derived from serial 10-fold dilutions of RNA.

METHODS. LC were enriched from BALB/c TE-EC (ref. 22), stained with M5/114.15.12 (anti-I-A^d and anti-I-E^d, ATCC) and FITC-goat (Fab')₂ anti-rat IgG (TAGO) and additionally purified using a BDIS Facstar Plus flow cytometer. Purity of sorted populations was analysed by repeat flow cytometry (*a*). Total RNA was extracted from sorted cells using RNAzol B (BIOEX) and precipitated with yeast transfer RNA. mRNAs were reverse-transcribed and amplified in serial 10-fold dilutions of RNA using a GeneAmp RNA PCR kit (Perkin Elmer Cetus) and primers (Synthecell) at 1 μM for both cDNA synthesis and amplification. The I-A α-chain primers used have been described²³. Sense (5'-AGACCAAGATCAATCCAC-3') and antisense (5'-CTCGCAACACCTTATGTAC-3') K1 primers correspond to a portion of the 3'-untranslated region of murine K1 mRNA sequence (S. Yuspa). Sense (5'-GTGAATCCCAAGAAGACCTGG-3') and anti-sense (5'-CTTCATCACGGA-GGTTCTG-3') E-cadherin primers encompassed the sequence encoding the epitopes reactive with ECCD-1 and ECCD-2 (refs 24, 25) and spanned three introns²⁶. PCR parameters included denaturation at 95 °C for 1 min, annealing at 55 °C for 1.15 min and polymerization at 72 °C for 2 min. Twenty-five cycles were used to amplify K1 mRNA and 30 cycles were used for I-A α-chain and E-cadherin mRNAs. PCR products (430 base pairs (bp) for I-A, 323 bp for K1 and 344 bp for E-cadherin) were resolved in 1.8% agarose (Gibco BRL) gels and detected after ethidium bromide staining. PCR products were not detected if RNA or reverse transcriptase was omitted from the reaction tubes.

endogenous synthesis, we characterized LC- and KC-derived E-cadherin messenger RNAs. LC were enriched to >98% purity (Fig. 2a) and total RNAs were prepared. LC- and KC-derived mRNAs were reverse-transcribed and complementary DNAs were amplified using the polymerase chain reaction (PCR) and E-cadherin, I-A α -chain, and keratin-1 (K1) primers. Reverse-transcribed PCR (RT-PCR) was done with serial dilutions of LC and KC RNA and the amounts of E-cadherin PCR products were compared with the amounts of I-A α -chain and K1 PCR products in each sample. E-cadherin PCR products of the predicted size were obtained from both LC and KC RNA samples (Fig. 2b). LC-derived samples were greatly enriched in I-A α -chain and greatly depleted of K1 mRNA when compared with amounts of KC RNA that gave similar amounts of E-cadherin PCR product. Thus, most of the E-cadherin PCR product derived from LC RNA originated in LC, and was not derived from contaminating KC.

Cadherins mediate cell-to-cell adhesion in tissues by binding to identical cadherins on adjacent (and usually ontogenically related) cells^{16,17}. To determine whether cadherins on LC could mediate intercellular adhesion through this type of homophilic interaction, we examined the ability of LC to bind to L cells (mouse fibroblasts devoid of cadherins) and L cells transfected with a vector encoding mouse E-cadherin (LEC cells) *in vitro*. As shown in Fig. 3, LEC cells expressed E-cadherin and avidly bound LC, whereas few LC attached to L cells lacking cadherins.

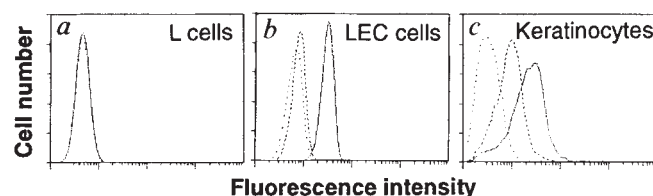


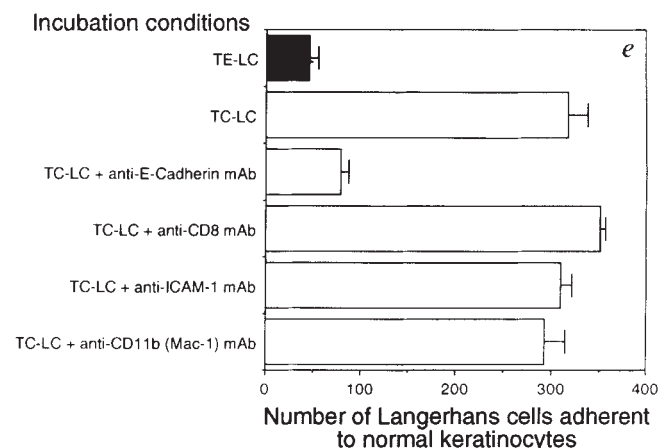
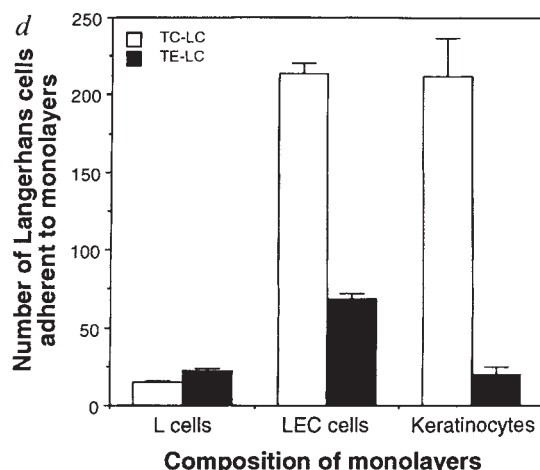
FIG. 3 Adhesion of Langerhans cells to E-cadherin-expressing fibroblasts and keratinocytes *in vitro*. L cells, E-cadherin-transfected L cells (LEC cells) and cultured murine KC were analysed for cadherin expression as described (Fig. 1). a, b, c. Profiles of L cells, LEC cells and KC stained with anti-CD8 (53-6.72, dotted line), anti-E-cadherin (ECCD-2, continuous line) or anti-P-cadherin (PCD-1, dashed line) antibody. The ability of TC-LC (open bars) and TE-LC (solid bars) to adhere to L cells, LEC cells and KC was tested (d). In a separate experiment, the influence of anti-E-cadherin (ECCD-1) and unrelated anti-adhesion molecule antibodies on LC-KC interactions was studied (e). mAb, Monoclonal antibody.

METHODS. L cells were transfected with expression vectors carrying the mouse E-cadherin cDNA (pBATEM2) and the neomycin-resistance gene (pBATneo^r) as described¹⁵. KC were prepared from neonatal BALB/c mice and cultured to confluence in low (0.05 mM)-calcium-containing medium²⁷. Low-calcium medium was replaced with normal calcium DMEM (Gibco BRL), 10% FBS and the incubation continued for 4 h (ref. 28) before initiating binding assays. Fibroblasts and KC (cultured in low-calcium medium) were detached from culture dishes after incubation in calcium-free HBSS, 2 mM EDTA and were washed in HBSS with calcium, stained for cadherins and analysed by flow cytometry. LC were enriched from TE- and TC-EC²², stained with M5/114.15.2, FITC-(Fab')₂ goat anti-rat IgG and analysed by flow cytometry. Preparations enriched in LC comprised 30–60% I-A⁺ cells. M5/114.15.2 was purified from supernatants using a protein G column (Pierce Chemical) and modified with NHS-LC-biotin (Pierce). Supernatants containing 53-6.72 (anti-CD8), YN/1.7.4 (anti-ICAM-1, ATCC) and M1/70/15/11/5/HL (anti-Mac-1 (CD11b), ATCC) were concentrated in Centrprep 30 microconcentrators (Amicon) before use in inhibition studies. PCD-1 (ref. 8) and ECCD-1 (ref. 18) were supplied by M. Takeichi. Antibody concentrations were assayed by indirect ELISA. To assess binding activity, LC stained with biotin-conjugated M5/114 and streptavidin-RED613 (Gibco BRL) were suspended in calcium-containing HBSS, 1% BSA, and allowed to adhere to confluent fibroblast or KC monolayers in 24-well culture plates (6×10^4 LC in 1 ml per well (d) or 5×10^4 LC in 0.2 ml per well (e)). In the experiment shown in e, wells contained anti-adhesion molecule antibodies ($100 \mu\text{g ml}^{-1}$) as indicated. After 1 h at 37 °C, wells were washed six times

Prior treatment of LC with trypsin in EDTA resulted in a 68% reduction in LC-LEC cell interactions. We also tested the ability of LC to adhere to murine KC expressing E- and P-cadherins (Fig. 3d). The number of TC-LC that adhered to KC was 10-fold greater than the number of TE-LC that bound, implying a role for cadherins in LC-KC interactions. By counting the number of LC in representative fluorescence photomicrographs of monolayers before and after removal of nonadherent LC, we determined that the majority ($57 \pm 13\%$, $n = 6$ experiments) of TC-LC added into the binding assay adhered to KC.

The importance of E-cadherin in LC-KC binding was confirmed by quantifying LC attachment to KC in the presence of antibody known to inhibit E-cadherin-dependent interactions¹⁸ and antibodies reactive with unrelated adhesion molecules. Addition of the anti-E-cadherin antibody ECCD-1 into the binding assay reduced adhesion of LC to KC by $81 \pm 4\%$ ($n = 3$ experiments); equal amounts of anti-CD8, anti-ICAM-1 and anti-Mac-1 antibody had no effect on LC binding (Fig. 3e).

We also assessed the level of cadherin expression on cultured LC (that may represent the *in vitro* equivalent of LC that have migrated to regional lymph nodes *in vivo*^{6,7}) and measured the ability of cultured LC to bind to KC, L cells and LEC cells. As compared with fresh TC-LC, TC-cultured LC expressed lower levels of E-cadherin (mean fluorescence intensity $34 \pm 8\%$ that of fresh TC-LC, $n = 3$ experiments) and adhered less readily to KC ($16 \pm 8\%$ as well as fresh TC-LC, $n = 3$ experiments)



with BSA, HBSS and adherent (fluorescent) LC were counted using an inverted phase fluorescence microscope. Data reported represent the mean $\pm$ s.e.m. of the number of fluorescent cells enumerated in triplicate wells (in five random fields encompassed by the $20\times$ objective (0.55 mm^2 per field) per well).

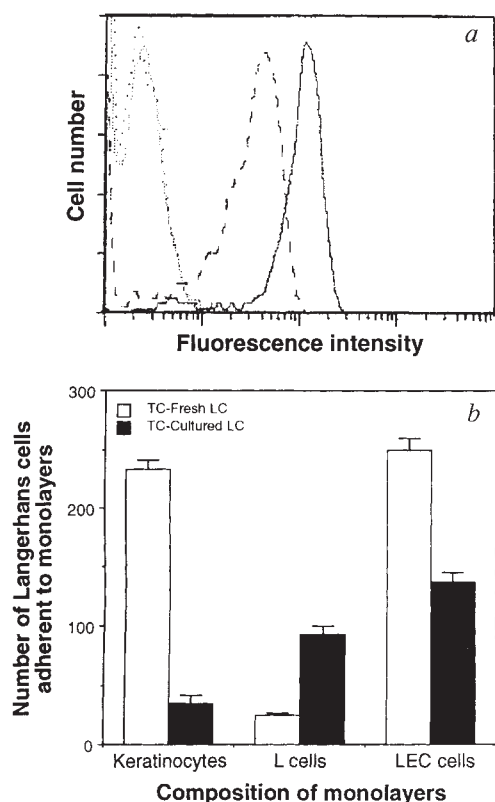


FIG. 4 Adhesion of fresh and cultured Langerhans cells to keratinocytes and fibroblasts *in vitro*. Fresh and cultured LC were stained for E-cadherin and analysed by flow cytometry (a). Profiles of TC-fresh LC stained with anti-CD8 (closely spaced dots) or anti-E-cadherin (ECCD-2, continuous line; mean fluorescence intensity (MFI)=114) and TC-cultured LC stained with anti-CD8 (widely spaced dots) or anti-E-cadherin (dashed line, MFI=33) are shown. The ability of TC-fresh LC and TC-cultured LC to attach to KC, L cells and LEC cells was also quantified (b).

METHODS. TC-fresh LC (39% I-A⁺) were prepared and analysed for cadherin expression as described in Fig. 1 legend. TC-cultured LC (45% I-A⁺) represent BALB/c EC that were cultured for 72 h (ref. 21), TC-treated and enriched by density gradient centrifugation²⁹. Cell-binding activity was measured as described for Fig. 3. Equal numbers of TC-fresh LC and TC-cultured LC (5×10^4 per well) were added into wells at the initiation of the binding assay.

(Fig. 4). The increased binding of TC-cultured LC to LEC cells relative to KC is largely reflective of the ability of TC-cultured LC to attach to fibroblasts, apparently through an E-cadherin-independent process (Fig. 4). These results further emphasize the potential importance of E-cadherin in LC-KC interactions *in vivo*.

Members of the cadherin family are prominently represented in epithelial and neural tissues, where they play important roles in morphogenesis and in maintenance of tissue integrity^{16,17}. Although structurally related molecules have recently been described in skeletal muscle¹⁹, cadherins are not ubiquitously expressed by mesenchymal cells and have not previously been identified on leukocytes. Our results indicate that epidermal LC express functional cadherins and suggest that cadherins may be responsible for certain biological properties of these highly specialized bone marrow-derived cells, including their propensity to persist in epidermis. It will be of interest to determine if cadherin expression is a general feature of leukocyte subpopulations that reside in epidermis, and if cadherins are involved in leukocyte-epithelial interactions in other tissues as well. □

- Frelinger, J. G., Hood, L., Hill, S. & Frelinger, J. A. *Nature* **282**, 321–323 (1979).
- Kripke, M. L., Dunn, C. G., Jeevan, A., Tang, J. & Bucana, C. J. *Immun.* **145**, 2833–2838 (1990).
- Schuler, G. & Steinman, R. M. *J. exp. Med.* **161**, 526–546 (1985).
- Romani, N. *et al. J. invest. Dermat.* **93**, 600–609 (1989).
- Nose, A. & Takeichi, M. *J. Cell Biol.* **103**, 2649–2658 (1986).
- Goodwin, L. *et al. Biochim. biophys. Acta* **173**, 1224–1230 (1990).
- Koch, P. J. *et al. Eur. J. Cell Biol.* **53**, 1–12 (1990).
- Collins, J. E. *et al. J. Cell Biol.* **113**, 381–391 (1991).
- Mechanic, S., Raynor, K., Hill, J. E. & Cowin, P. *Proc. natn. Acad. Sci. U.S.A.* **88**, 4476–4480 (1991).
- Amagai, M., Klaus-Kovtun, V. & Stanley, J. R. *Cell* **67**, 869–877 (1991).
- Shirayoshi, Y., Nose, A., Iwasaki, K. & Takeichi, M. *Cell Struct. Funct.* **11**, 245–252 (1986).
- Nose, A., Nagafuchi, A. & Takeichi, M. *Cell* **54**, 993–1001 (1988).
- Takeichi, M. *A. Rev. Biochem.* **59**, 237–252 (1990).
- Takeichi, M. *Science* **251**, 1451–1455 (1991).
- Yoshida-Noro, C., Suzuki, N. & Takeichi, M. *Dev. Biol.* **101**, 19–27 (1984).
- Donalies, M., Cramer, M., Ringwald, M. & Starzinski-Powitz, A. *Proc. natn. Acad. Sci. U.S.A.* **88**, 8024–8028 (1991).
- Hauser, C., Snapper, C. M., Ohara, J., Paul, W. E. & Katz, S. I. *Eur. J. Immun.* **19**, 245–251 (1989).
- Tang, A. & Udey, M. C. *J. Immun.* **146**, 3347–3355 (1991).
- Rasanen, L., Lehto, M., Reunala, T. & Leinikki, P. *J. invest. Dermat.* **86**, 9–12 (1986).
- Witsell, A. B. & Schook, L. B. *Biotechniques* **9**, 318–322 (1990).
- Nagafuchi, A., Shirayoshi, Y., Okazaki, K., Yasuda, K. & Takeichi, M. *Nature* **329**, 341–343 (1987).
- Nose, A., Tsuji, K. & Takeichi, M. *Cell* **61**, 147–155 (1990).
- Ringwald, M., Baribault, H., Schmidt, C. & Kemler, R. *Nucleic Acid Res.* **19**, 6533–6539 (1991).
- Yuspa, S. H. in *Methods in Skin Research* (eds Skerrow, D. & Skerrow, C. J.) 213–249 (Wiley, Chichester, 1985).
- Hennings, H. *et al. Cell* **19**, 245–254 (1980).
- Tang, A. & Udey, M. C. *Eur. J. Immun.* **22**, 581–586 (1992).

ACKNOWLEDGEMENTS. We thank M. Takeichi (Kyoto University) for antibodies ECCD-1, ECCD-2 and PCD-1 and expression vectors pBATEM2 and pBATEOβ. S. Yuspa (NCI) for unpublished murine K1 sequence information and assistance in culturing normal keratinocytes. S. Sharrow (NCI) for advice on flow cytometry. V. McFarland and R. O'Neal for technical assistance. H. Schaefer for preparing the figures and A. Erik, S. Katz, S. Shaw and J. Vogel for reading the manuscript.

Metal ion catalysis in the *Tetrahymena* ribozyme reaction

Joseph A. Piccirilli, Joseph S. Vyle*, Marvin H. Caruthers & Thomas R. Cech†

Howard Hughes Medical Institute, and Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309, USA

ALL catalytic RNAs (ribozymes) require or are stimulated by divalent metal ions, but it has been difficult to separate the contribution of these metal ions to formation of the RNA tertiary structure¹ from a more direct role in catalysis. The *Tetrahymena* ribozyme catalyses cleavage of exogenous RNA^{2,3} or DNA^{4,5} substrates with an absolute requirement for Mg²⁺ or Mn²⁺ (ref. 6). A DNA substrate, in which the bridging 3' oxygen atom at the cleavage site is replaced by sulphur, is cleaved by the ribozyme about 1,000 times more slowly than the corresponding unmodified DNA substrate when Mg²⁺ is present as the only divalent metal ion. But addition of Mn²⁺ or Zn²⁺ to the reaction relieves this negative effect, with the 3' S–P bond being cleaved nearly as fast as the 3' O–P bond. Considering that Mn²⁺ and Zn²⁺ coordinate sulphur more strongly than Mg²⁺ does^{7,8}, these results indicate that the metal ion contributes directly to catalysis by coordination to the 3' oxygen atom in the transition state, presumably stabilizing the developing negative charge on the leaving group. We conclude that the *Tetrahymena* ribozyme is a metalloenzyme, with mechanistic similarities to several protein enzymes^{9–12}.

Shortened forms of the self-splicing intervening sequence RNA of *Tetrahymena thermophila* catalyse a sequence-specific endonuclease reaction^{2,3}. It is analogous to the first step in splicing of pre-ribosomal RNA but occurs with multiple turnover (Fig. 1a). The substrate, which can be RNA or DNA^{4,5}, binds by base-pairing to an internal guide sequence on the ribozyme^{2,13}. Guanosine (G) binds at a saturable binding site^{14,15} located within the folded core of the ribozyme¹⁶, and the 3' hydroxyl of guanosine attacks the reaction-site phosphoryl group^{17,18}. Water or OH[−] can replace guanosine in a sequence-

Received 2 July; accepted 21 October 1992.

- Stingl, G., Katz, S. I., Clement, L., Green, I. & Shevach, E. M. *J. Immun.* **121**, 2005–2013 (1978).
- Breathnach, S. M. *Epidermal Langerhans Cells* (ed. Schuler, G.) 23–47 (CRC, Boca Raton, 1991).
- Katz, S. I., Tamaki, K. & Sachs, D. H. *Nature* **282**, 324–326 (1979).

* Present address: The Robert Robinson Laboratories, Department of Chemistry, University of Liverpool, Liverpool L69 3BX, UK.

† To whom correspondence should be addressed.

specific hydrolysis reaction¹⁵. The rate constants and equilibrium binding constants for the RNA and DNA reaction pathways have been described^{4,15,19–21}. For the DNA substrate, unlike the RNA substrate, the chemical cleavage step is rate-limiting at all concentrations of guanosine and substrate (10 mM MgCl₂; 30 or 50 °C).

We synthesized²² 3'S DNA substrates, which have the bridging 3' oxygen atom of the reactive phosphoryl group replaced by sulphur (Fig. 1*b*). Phosphorothioate diesters in which one of the non-bridging oxygens is replaced by sulphur have been frequently used as mechanistic probes in enzymology^{23–25}, including studies with this ribozyme^{17,18,26}. In contrast, the 3'S derivative has only recently become available²². When the 5'³²P-end-labelled substrates and L-21 *ScaI* ribozyme³ were incubated with 12 mM Mg²⁺ as the divalent metal ion, only the unmodified substrate reacted well (Fig. 2*a*). The 5'-³²P-labelled 3'S substrate d(CCCUCU_sA) was cleaved extremely slowly, with a rate constant for the chemical cleavage step ($k_c^{3'S}$) ~10³-fold smaller than that of the oxy-containing substrate, $k_c^{3'O}$ (Table

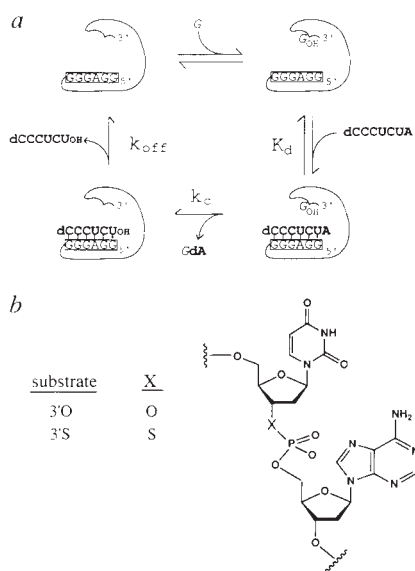


FIG. 1 *a*, Endonuclease reaction of the L-21 *ScaI* ribozyme. The ribozyme is a shortened version of the *Tetrahymena* pre-rRNA intervening sequence missing 21 nucleotides from the 5' end and 3–4 nucleotides from the 3' end. The ribozyme binds an oligonucleotide substrate, in this case the single-stranded DNA substrate d(CCCUCU_sA), by base pairing to the internal guide sequence (IGS) (boxed)^{2,13}. Although G is shown binding before the DNA substrate, there is no compulsory order to binding¹⁵. *b*, Oligodeoxyribonucleotides synthesized for this work. Unmodified substrates containing normal phosphodiester linkages are designated as 3'O. Modified substrates in which the 3' oxygen atom of the dU at the cleavage site is replaced by sulphur are indicated by d(U_sA) at the position of the phosphorothioate linkage and designated as 3'S substrates. The product of the reaction of the 3'S substrate is a 3' sulphhydryl rather than the normal 3' hydroxyl.

METHODS. Preparation of DNA with the 3'-S-thiophosphoramidate of dU was as described^{22,28} using tetrabutylammonium periodate as oxidizing agent. Oligonucleotides were purified before removal of the 5'-dimethoxytrityl group by reverse-phase HPLC, detritylated with 80% acetic acid, and purified by gel electrophoresis. 2'-Deoxy-3'-thiouridine was prepared as the 5'-monomethoxytrityl derivative by a procedure analogous to the preparation of 2'-deoxy-3'-thiothymidine²², starting from 2'-deoxyuridine. 5'-O-monomethoxytrityl-3'-thiouridine was characterized by ¹H-NMR (300 MHz, in CDCl₃, using tetramethylsilane as internal standard): δ 2.40 (2H, m, H2' and H2''), 3.50 (2H, m, H5' and H5''), 3.75 (4H, m, H3', CH₃O), 5.24 (1H, m, H4'), 6.04 (1H, m, H1'), 6.80 (2H, d, ArH), 7.20–7.36 (12H, m, ArH, H5), 8.04 (1H, d, H6), 9.13 (1H, s, NH, exchangeable). Phosphitylation of 5'-monomethoxy-2'-deoxy-3'-thiouridine using 2-cyanoethyl *N,N*-diisopropylaminochlorophosphoramidate under standard conditions produced the 3'-S-phosphorothioamidate derivative in a yield of 70%. ³¹P-NMR (90 MHz, in CDCl₃, using 85% phosphoric acid as external standard) gave δ 159.94, 164.99.

1*a*). Increasing the Mg²⁺ concentration to 200 mM increased the individual rate constants, but the 3'S substrate still reacted 10³-fold more slowly than the 3'O substrate. The low level of reaction of the modified substrate was not due to its failure to bind to the ribozyme, because direct measurement of K_d by non-denaturing gel electrophoresis²⁷ indicated that the 3'S substrate actually bound 2–4-fold more tightly than the unmodified 3'O substrate (Table 1*a*). This is consistent with the previous observations that a central 3'S phosphorothioate linkage does not significantly destabilize an oligonucleotide duplex²⁸. The observation of a large thio effect associated with the chemical step of the ribozyme reaction was unexpected because the intrinsic chemical rate of cleavage of the sulphur-phosphorus

TABLE 1 Kinetic parameters for ribozyme cleavage of DNA and thio DNA substrates with different divalent metal ions

(a)				
M ²⁺ (mM)	Substrate 5'- ³² P- d(CCCUCU _s A)	k_c^* (10 ⁻⁴ min ⁻¹)	$K_m^\dagger$ or $K_d^\ddagger$ (nM)	$k_c(-G)^\S$ (10 ⁻⁴ min ⁻¹)
Mg ²⁺ (12)	3'O	8.9	270 $\ddagger$	0.068
	3'S	0.009	100 $\ddagger$	ND $\parallel$
Mg ²⁺ (200)	3'O	99		
	3'S	0.099		
Mg ²⁺ (10); Mn ²⁺ (2)	3'O	29	~41 $\ddagger$	0.33
	3'S	3.7	~24 $\ddagger$	0.12
Mn ²⁺ (12)	3'O	21		
	3'S	6.0		
(b)				
	d(CCCUCU _s - AAAAA)- ³² P3'dA	$k_{obs}^\P$ (10 ⁻⁴ min ⁻¹)		
Mg ²⁺ (12)	3'O	2.4		
	3'S	<0.002		
Mg ²⁺ (10); Mn ²⁺ (2)	3'O	23		
	3'S	6.4		
Mg ²⁺ (10); Zn ²⁺ (2)	3'O	190		
	3'S	24		
Mg ²⁺ (10); Cd ²⁺ (2)	3'O	<0.038		
	3'S	0.089		
Mg ²⁺ (10); Co ²⁺ (2)	3'O	2.8		
	3'S	0.05		

Reactions were as described in Fig. 2 using the M²⁺ conditions indicated here. Previous characterization of DNA cleavage by the ribozyme was at 50 °C and 10 mM MgCl₂ (ref. 4). To minimize the extent of metal-catalysed hydrolysis of the ribozyme, which was problematic at 50 °C with Zn²⁺ and Mn²⁺, reactions were done at 30 °C. Reactions with 10 mM Mg²⁺ and 2 mM Mn²⁺ were linear for up to 72 h. Even the reactions with Mn²⁺ alone proceeded well up to 30 h. Reactions with 10 mM Mg²⁺ and 2 mM Zn²⁺ were linear up to ~3 h. Radioactivity in substrate and product was measured by direct scanning of the dried gel (Phosphorimager, Molecular Dynamics). Observed rate constants were determined from the slope of semilogarithmic plots of the fraction of substrate remaining (S_t/S_0 , where S_t is the amount of substrate at time t) versus t , or from the slope of initial rate plots of $\%P = P_t/(P_t + S_t) \times 100$, where P_t is the amount of product at time t . Plots were linear over the first 8% of product formed. Variations in rate constants of <20% were observed for reactions run side by side. But variations as large as 2-fold are observed for reactions on different days, especially for the reactions containing Mn²⁺.

* Rate constant for the chemical conversion of the E-G-S ternary complex (ref. 4; Fig. 1*a*) obtained by measuring k_{obs} at saturating ribozyme concentration (3.3 μM) and saturating G (2.5 mM).

† Determined from plots of k_{obs} versus ribozyme concentration. Ribozyme concentration was varied from 5 nM to 300 nM.

‡ Binding was measured directly using a gel-mobility shift method²⁷. The electrophoresis buffer was 34.5 mM Tris, 65.5 mM HEPES, 0.1 mM disodium-EDTA and 12 mM MgCl₂, pH 7.5, and the gel was maintained at 30 °C. Ribozyme and substrate were equilibrated for 2 h at 30 °C.

§ Rate constant for the hydrolysis reaction, representing the chemical conversion of the E-S complex in the absence of guanosine.

|| Not detected.

¶ Observed rate constant under conditions described for Fig. 2*b*.

bond is predicted to be faster than that of the normal phosphate diester^{29,30}.

With 2 mM Mn^{2+} and 10 mM Mg^{2+} present in the reaction, the thio substrate was readily cleaved (Fig. 2a); $k_c^{3'S}$ was now only ~8-fold smaller than $k_c^{3'O}$ (Table 1a). Although addition of Mn^{2+} also accelerated the rate of cleavage of the 3'O substrate, the acceleration of the 3'S substrate by Mn^{2+} was greater by a factor of more than 10^2 . The hydrolysis reaction of the 3'S substrate also shows a requirement for Mn^{2+} ($k_c(-G)$; Table 1a). The specificity of the thio substrate for Mn^{2+} is consistent with the ability of Mn^{2+} to coordinate sulphur more strongly than Mg^{2+} (refs 7 and 8). When Mg^{2+} is replaced by Mn^{2+} as the only divalent metal ion, the thio effect for ribozyme cleavage is relieved by a factor of ~280 ((8.9/0.009)/(21/6); Table 1a), similar to the ratio of the relative binding of these two metal ions to the O and S atoms of ATP- β S (31,000 for Mg^{2+} /~170 for Mn^{2+} ~200; ref. 8). This almost quantitative agreement supports the conclusion that the relative thio effect in Mg^{2+} versus Mn^{2+} reflects the differential ability of these metal ions to coordinate oxygen versus sulphur in the transition state.

Measurement of K_m (which should equal K_d ; ref. 4) for reaction in 2 mM Mn^{2+} plus 10 mM Mg^{2+} showed that binding

of both the 3'O and 3'S substrates was tighter than in 12 mM Mg^{2+} . However, the 3'S substrate bound about 2-fold more tightly than the 3'O substrate, regardless of the presence of Mn^{2+} (Table 1a). Thus, the interaction of the metal ion with the 3' oxygen or 3' sulphur atom appears insignificant in the ground-state E·S·G complex (E is ribozyme, S is substrate; Fig. 1a).

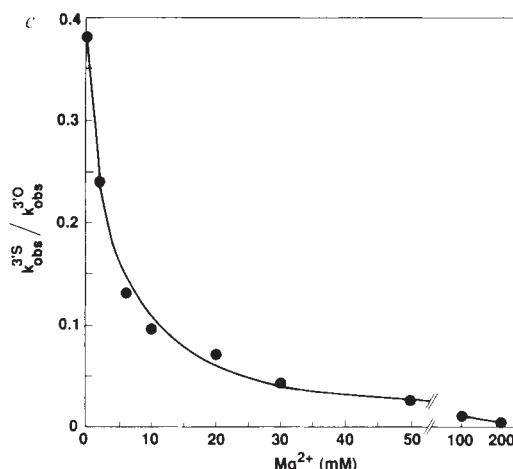
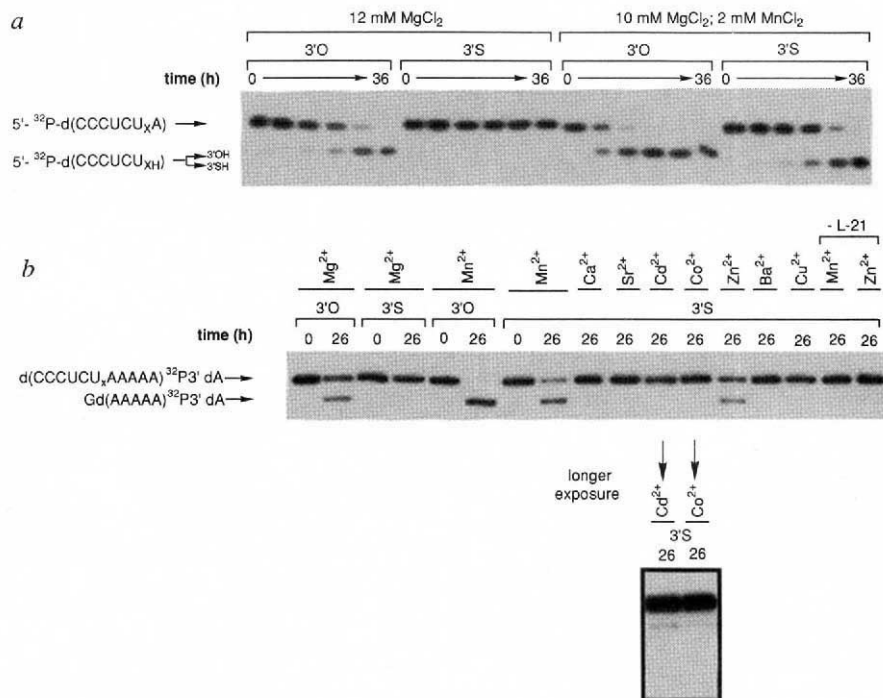
Of all the divalent cations screened, only Mn^{2+} and Zn^{2+} were able to restore cleavage of the 3'S substrate to rates comparable to that of the unmodified substrate (Fig. 2b; Table 1b). Low levels of cleavage were observed in the presence of Cd^{2+} and Co^{2+} (Fig. 2b, inset). Co^{2+} had little effect on the reaction rate of the 3'O substrate, but Cd^{2+} strongly inhibited the 3'O reaction such that the 3'S substrate actually reacted faster (Table 1b). Because Zn^{2+} , Cd^{2+} and Co^{2+} are also metal ions that can coordinate sulphur effectively^{7,8}, their ability to restore cleavage of the thio substrate supports the interpretation of direct metal ion coordination.

To test whether the Mn^{2+} site responsible for the reaction of the 3'S substrate is also a Mg^{2+} site, we did a competition experiment. Increasing the concentration of Mg^{2+} relative to Mn^{2+} inhibited the reaction of the 3'S substrate relative to that of the unmodified substrate (Fig. 2c). Thus, the data are most

FIG. 2 Metal ion specificity switch in cleavage of 3'S substrate.

a, Endonuclease reaction of 5'-³²P-labelled substrates 3'O (5'-³²P-d(CCCUCUA)) and 3'S (5'-³²P-d(CCCUCU_SA)) for 0, 2, 4, 8, 18 and 36 h. Cleavage of the 3'S substrate by the ribozyme forms a product 5'-³²P-d(CCCUCU_SH) containing a sulphhydryl group that has a greater mobility than the product 5'-³²P-d(CCCUCU_OH) under these electrophoresis conditions. To prevent oxidation of the product, 5 mM dithiothreitol (DTT) was included in the reaction; 10 mM DTT was present in the gel and in the electrophoresis buffer. Other conditions were as described below. Gels were prerun for 6 h before loading. If DTT was excluded from this procedure, 5'-³²P-d(CCCUCU_SH) ran as a smear rather than as a single band, making quantitation difficult. b, Endonuclease reaction of 3'-end-labelled DNA substrates 3'O [d(CCCUCUAAAA)³²P3'dA] and 3'S [d(CCCUCU_SAAAA)³²P3'dA] in the presence of Mg^{2+} and other divalent metal ions. Phosphotransferase with G results in a faster-migrating product Gd(AAAAA)³²P3'dA which can be electrophoretically separated from the product that would arise from metal-catalysed hydrolysis of the sulphur-phosphorus bond off the ribozyme (5'-P-d(AAAAA)³²P3'dA). The inset shows a longer exposure of the Co^{2+} and Cd^{2+} reactions to allow visualization of the product. The last two lanes represent incubation of substrate in buffer and metal ion in the absence of ribozyme and show that degradation of substrate is undetectable in the absence of ribozyme. c, Evidence that Mn^{2+} and Mg^{2+} compete for the same binding site. The value of k_{obs} was determined as described in Table 1.

METHODS. Ribozyme was renatured by incubation for 30 min at 50 °C in 50 mM MES, pH 7.0, 10 or 12 mM $MgCl_2$ and 2 mM guanosine, followed by incubation for 5 min at 30 °C before addition of DNA substrate. Reaction conditions were 30 °C, 1 μ M L-21 *Scal* ribozyme, ~1 nM 3'O or 3'S substrate, 50 mM MES, pH 7.0, 2.5 mM G, and 12 mM $MgCl_2$, or 10 mM $MgCl_2$ and 2 mM of another divalent metal ion as indicated (for c, 1 mM $MnCl_2$ plus $MgCl_2$ as indicated). Reactions were quenched at 0 °C by addition to twice the volume of 20 mM EDTA, 10 M urea, 0.04% xylene cyanol, 1 mM Tris-HCl, pH 7.5, and electrophoresed on a 20% polyacrylamide/7 M urea gel. Metal ion chlorides (Aldrich) were all $\geq 99.99\%$ pure.



easily interpreted in terms of a direct interaction between a ribozyme-bound metal ion and the leaving oxyanion or thioanion in the transition state (Fig. 3). For the oxy-containing substrate, the catalytic metal can be Mg^{2+} or Mn^{2+} , as both are effective at stabilizing the developing negative charge on the oxyanion leaving group. For the thio-containing substrate, Mg^{2+} still binds but is ineffective at stabilizing the thioanion leaving group, and Mn^{2+} or Zn^{2+} must be present for efficient catalysis. Studies on non-enzymatic reactions of phosphate monoesters³¹, and more recently phosphate diesters³², suggest that divalent metal ions can stabilize an oxyanion leaving group.

The hammerhead ribozyme also employs Mg^{2+} or Mn^{2+} in active-site chemistry, in this case coordinating with the *pro*-R_P oxygen of the reaction-site phosphoryl group and presumably serving to stabilize the pentacoordinate transition state^{33,34}. Interaction of the leaving group or nucleophile with a metal ion in the hammerhead reaction has been proposed^{33,35} but not tested directly. Perhaps ribozymes will in general use metal ion catalysis for stabilization of leaving groups and activation of nucleophiles instead of using functional groups on the RNA as general acid-base catalysts (ref. 36; compare ref. 37).

It has been proposed that a number of protein phosphotransferases use a two-magnesium-ion mechanism and that such a mechanism would be well suited to ribozyme catalysis^{9,10}. The metal specificity switch described in this work supports the

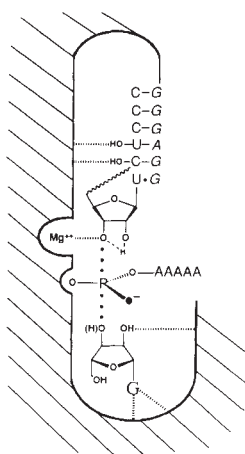


FIG. 3 Postulated transition-state interactions of RNA substrate and ribozyme in the endonuclease reaction, including the metal ion interaction (asterisk) inferred from this work. We have identified one ligand for this catalytic Mg^{2+} in the transition state. Given the preference for Mg^{2+} to be 6- or 7-coordinate³⁸, the ribozyme presumably provides the appropriate spatial arrangement of heteroatomic ligands to bind and orient the metal in a manner that will facilitate catalysis (E. Christian and M. Yarus, manuscript submitted). The three-dimensional ribozyme surface is represented by the hatched pattern outside the dark outline. The unhatched interior represents the ribozyme active site. The RNA substrate is bound to the ribozyme by base pairing to the IGS^{2,13} (represented by upper-case italic letters) and interactions with specific 2'-hydroxyl groups^{20,21,39}. Guanosine is bound by at least three different interactions^{14,16,40}. Nucleophilic attack at the reaction-site phosphoryl group by the 3'-hydroxyl of guanosine is an in-line, S_N2 (P) reaction, as inferred by the inversion of stereochemistry at phosphorus^{17,18}. Substitution of the *pro*-R_P oxygen atom (●) at the cleavage site with sulphur gives a small thio effect consistent with that expected from the intrinsic chemistry, suggesting that there is no specific interaction with this atom in the transition state²⁶. Substitution of the *pro*-S_P oxygen atom with sulphur, on the other hand, reduces activity >1,000 fold, suggesting a direct contact with the ribozyme or ribozyme-bound Mg^{2+} ion (shown by the pocket in the dark outline). But this thio effect cannot be relieved by addition of Mn^{2+} to the reaction (J.A.P. and T.R.C., manuscript in preparation). The hydrogen-bond donation from the 2'-hydroxyl group to the 3' oxygen in the transition state was inferred from a study of the rate of the chemical step of the reaction with different 2' substituents (D. Herschlag, F. Eckstein and T.R.C., manuscript in preparation). Dashed lines, H bonds or metal ion-oxygen coordination. Dotted P-O bonds, bonds partially formed or partially broken in the transition state.

proposed role of one of the magnesium ions in leaving-group stabilization. The possibility that a second metal ion activates the 3' hydroxyl of G for nucleophilic attack has not been tested. Nevertheless, at least some aspects of metal ion catalysis appear to be common to the *Tetrahymena* ribozyme and certain protein enzymes. □

Received 2 September; accepted 13 November 1992.

1. Celander, D. W. & Cech, T. R. *Science* **25**, 401-407 (1991).
2. Zaug, A. J., Been, M. D. & Cech, T. R. *Nature* **324**, 429-433 (1986).
3. Zaug, A. J., Grosshans, C. A. & Cech, T. R. *Biochemistry* **27**, 8924-8930 (1988).
4. Herschlag, D. & Cech, T. R. *Nature* **344**, 405-409 (1990).
5. Robertson, D. L. & Joyce, G. F. *Nature* **344**, 467-468 (1990).
6. Grosshans, C. A. & Cech, T. R. *Biochemistry* **28**, 6888-6894 (1989).
7. Jaffe, E. K. & Cohn, M. *J. biol. Chem.* **253**, 4823-4825 (1978).
8. Pecoraro, V. L., Hermes, J. D. & Cleland, W. W. *Biochemistry* **23**, 5262-5271 (1984).
9. Freemont, P. S., Friedman, J. M., Beese, L. S., Sanderson, M. R. & Steitz, T. A. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8924-8928 (1988).
10. Beese, L. S. & Steitz, T. A. *EMBO J.* **10**, 25-33 (1991).
11. Kim, E. E. & Wyckoff, H. W. *J. molec. Biol.* **218**, 449-464 (1991).
12. Hough, E. et al. *Nature* **338**, 357-360 (1989).
13. Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A. & Sczacchio, C. *Nature* **300**, 719-724 (1982).
14. Bass, B. L. & Cech, T. R. *Nature* **308**, 820-826 (1984).
15. Herschlag, D. & Cech, T. R. *Biochemistry* **29**, 10172-10180 (1990).
16. Michel, F., Hanna, M., Green, R., Bartel, D. P. & Szostak, J. W. *Nature* **342**, 391-395 (1989).
17. McSwiggen, J. A. & Cech, T. R. *Science* **244**, 679-683 (1989).
18. Rajagopal, J., Doudna, J. A. & Szostak, J. W. *Science* **244**, 692-694 (1989).
19. Herschlag, D. & Cech, T. R. *Biochemistry* **29**, 10159-10171 (1990).
20. Pyle, A. M. & Cech, T. R. *Nature* **350**, 628-631 (1991).
21. Bevilacqua, P. C. & Turner, D. H. *Biochemistry* **30**, 10632-10640 (1991).
22. Cosstick, R. & Vyle, J. S. *Nucleic Acids Res.* **8**, 2295-2300 (1990).
23. Knowles, J. R. A. *Rev. Biochem.* **49**, 877-919 (1980).
24. Eckstein, F. A. *Rev. Biochem.* **54**, 367-402 (1985).
25. Frey, P. A. *Adv. Enzym. rel. Areas molec. Biol.* **62**, 119-201 (1989).
26. Herschlag, D., Piccirilli, J. A. & Cech, T. R. *Biochemistry* **30**, 4844-4854 (1991).
27. Pyle, A. M., McSwiggen, J. A. & Cech, T. R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 8187-8191 (1990).
28. Vyle, J. S., Connolly, B. A., Kemp, D. & Cosstick, R. *Biochemistry* **31**, 3012-3018 (1992).
29. Milstien, S. & Fife, T. H. *J. Am. chem. Soc.* **89**, 5820-5826 (1967).
30. Nakamaye, K. L., Gish, G., Eckstein, F. & Vosberg, H.-P. *Nucleic Acids Res.* **16**, 9947-9958 (1988).
31. Herschlag, D. & Jencks, W. P. *J. Am. chem. Soc.* **109**, 4665-4674 (1987).
32. Browne, K. A. & Bruice, T. C. *J. Am. chem. Soc.* **114**, 4951-4958 (1992).
33. Dahm, S. & Uhlenbeck, O. C. *Biochemistry* **30**, 9464-9469 (1991).
34. Koizumi, M. & Ohtsuka, E. *Biochemistry* **30**, 5145-5150 (1991).
35. Perreault, J.-P., Labuda, D., Usman, N., Yang, J.-H. & Cedergren, R. *Biochemistry* **30**, 4020-4025 (1991).
36. Cech, T. R., Herschlag, D., Piccirilli, J. A. & Pyle, A. M. *J. biol. Chem.* **267**, 17479-17482 (1992).
37. Guerrier-Takada, C., Haydock, K., Allen, L. & Altman, S. *Biochemistry* **25**, 1509-1515 (1986).
38. Stezowski, J. J., Countryman, R. & Hoard, J. L. *Inorg. Chem.* **12**, 1749-1754 (1973).
39. Pyle, A. M., Murphy, F. L. & Cech, T. R. *Nature* **358**, 123-128 (1992).
40. Yarus, M., Illangsekere, M. & Christian, E. *J. molec. Biol.* **222**, 995-1012 (1991).

ACKNOWLEDGEMENTS. T.R.C. is an Investigator of the Howard Hughes Medical Institute and an American Cancer Society Professor. This research was supported in part by a grant from the NIH to M.H.C. We thank the W. M. Keck Foundation for support of RNA science on the Boulder Campus.

Solution structure of the cyclosporin A/cyclophilin complex by NMR

Yves Thériault, Timothy M. Logan, Robert Meadows, Liping Yu, Edward T. Olejniczak, Thomas F. Holzman, Robert L. Simmer & Stephen W. Fesik*

Pharmaceutical Discovery Division, Abbott Laboratories, Abbott Park, Illinois 60064, USA

CYCLOSPORIN A, a cyclic undecapeptide, is a potent immunosuppressant that binds to a peptidyl-prolyl *cis-trans* isomerase^{1,2} of 165 amino acids, cyclophilin³. The cyclosporin A/cyclophilin complex inhibits the calcium- and calmodulin-dependent phosphatase, calcineurin⁴, resulting in a failure to activate genes encoding interleukin-2 and other lymphokines^{5,6}. The three-dimensional structures of uncomplexed cyclophilin⁷, a tetrapeptide/cyclophilin complex^{8,9}, and cyclosporin A when bound to cyclophilin^{10,11} have been reported. However, the structure of the cyclosporin A/cyclophilin complex has not been determined. Here we present the solution structure of the cyclosporin A/cyclophilin complex

* To whom correspondence should be addressed.

obtained by heteronuclear three-dimensional NMR spectroscopy. The structure, one of the largest determined by NMR, differs from proposed models of the complex¹²⁻¹⁴ and is analysed in terms of the binding interactions and structure/activity relationships for CsA analogues^{15,16}.

Figure 1 depicts a stereoview of 31 superimposed structures of the cyclosporin A (CsA)/cyclophilin (CyP) complex generated from a total of 2,115 NMR-derived restraints using a distance geometry/simulated annealing protocol¹⁷. As shown in Table 1, the structures satisfy the distance restraints with no violation greater than 0.4 Å and show good covalent geometry and non-bonded contacts. For most of the complex, the structure is well defined by the NMR data. The atomic root-mean-square (r.m.s.) distribution about the mean coordinate positions for residues 5-165 of the CsA/CyP complex is 0.93 ± 0.10 Å for the backbone atoms and 1.51 ± 0.10 Å for all atoms. The least defined regions are residues 1-4 at the N terminus of CyP and the loop composed of CyP residues 41-45.

As shown in Fig. 2, the structure of CyP with bound CsA consists of an eight-stranded antiparallel β -sheet, two α -helices, and a series of loops and turns, as previously observed for uncomplexed CyP^{7,18} and a tetrapeptide/CyP complex^{8,9}. In general, the backbone of free CyP⁷ lies within the ensemble of NMR-derived solution structures for the complex (Fig. 1). Thus, unlike other proteins that undergo large structural changes upon ligand binding (such as penicillopepsin)¹⁹, the structure of cyclophilin changes very little after complexing with CsA. In contrast, the conformation of CsA when bound to cyclophilin^{10,11} is very different from either the crystal²⁰ or chloroform-solution structures^{20,21} of uncomplexed CsA. Although the bound conformation of CsA may pre-exist in aqueous solution²² to a small extent, NMR studies of a water-soluble CsA derivative ([D-diaminobutyric acid]CsA; ref. 8) indicate the presence of several different slowly interconverting conformations in aqueous solution (unpublished observations). Some of these forms may need to change conformation to allow binding to CyP, consistent with recent kinetic evidence that support a time-dependent inhibition of the peptidyl-prolyl *cis-trans* isomerase activity of cyclophilin by CsA²³.

FIG. 1 Superposition of 31 CsA/CyP NMR-derived structures (blue) and the X-ray structure of uncomplexed CyP (yellow)⁷ onto the mean coordinates for CsA/CyP. CsA is colour-coded in red. These structures were generated with a distance geometry/simulated annealing (DG/SA) protocol¹⁷ using the XPLOR/DG program²⁸. Structure calculations used 1,699 CyP/CyP, 227 CsA/CsA and 81 CsA/CyP proton-proton distance restraints obtained from heteronuclear three-dimensional nuclear Overhauser effect (NOE) spectra of CsA/[U-¹³C, ¹⁵N]CyP, CsA/[U-¹⁵N]CyP or [U-¹³C]CsA/CyP. The NOEs were interpreted using previously reported CyP backbone²⁹ and CsA assignments¹¹ and recently obtained CyP side-chain assignments, which included the stereospecific assignments of the Val and Leu methyl groups of CyP by biosynthetic fractional ¹³C-labelling³⁰. NOE-derived distance restraints were classified as 1.8-3.0, 1.8-4.0, 1.8-5.0 and 1.8-6.0 Å on the basis of the integrated NOE cross peak volumes. In addition to the NOE-derived distance restraints, 108 hydrogen bonds were included in the final structure calculations based on the identification of the slowly exchanging amide protons from a series of ¹H/¹⁵N correlation spectra of the CsA/[U-¹⁵N]CyP complex in ²H₂O and by visual inspection of preliminary structures of the complex derived solely from the NOE data. Calculations began by generating starting cartesian coordinates for the β and γ atoms from the covalent geometry and experimental restraints. The remaining coordinates were obtained by least-squares fitting the side chains to the backbone positions before the next step of the refinement procedure, which consisted of high-temperature (2,000 K) molecular dynamics followed by dynamical simulated annealing¹⁷. During this stage quadratic harmonic potentials were included for the covalent geometry, a square-well quadratic term was used for the distances and dihedral restraints, and a quartic

TABLE 1 Structural statistics and RMS deviations for 31 CsA/CyP structures*

Structural statistics	(SA)	($\overline{SA}$) _r
R.m.s.d. (Å) from experimental distance restraints†		
All (2,115)	0.019 ± 0.001	0.017
Inter-residue long (575)	0.017 ± 0.003	0.015
Inter-residue short (223)	0.024 ± 0.004	0.028
Inter-residue sequential (475)	0.020 ± 0.005	0.016
Intraresidue (426)	0.012 ± 0.002	0.009
Intraligand (227)	0.023 ± 0.001	0.024
Intermolecular (81)	0.010 ± 0.005	0.012
Hydrogen bonds (108)	0.020 ± 0.007	0.010
XPLOR potential energies‡		
F_{TOTAL} (kcal mol ⁻¹)	-4.0 ± 50.5	-29.7
F_{NOE} (kcal mol ⁻¹)	38.3 ± 5.8	32.4
$E_{\text{L-J}}$ (kcal mol ⁻¹)	-291.3 ± 31.9	-298.8
R.m.s.d. from idealized geometry used within XPLOR		
Bonds (Å)	0.00 ± 0.00	0.00
Angles (deg)	0.45 ± 0.02	0.44
Impropers (deg)	0.41 ± 0.02	0.41
Cartesian coordinate R.m.s.d. (Å)	Backbone	All non-H
(SA) versus ($\overline{SA}$) all§	0.93 ± 0.10	1.51 ± 0.10
(SA) versus X-Ray	1.47 ± 0.10	2.30 ± 0.15
($\overline{SA}$) _r versus X-Ray¶	1.21	1.98

* Where (SA) are the 31 NMR-derived structures, ($\overline{SA}$) is the mean atomic structure obtained by averaging the least squares fitted coordinates for the 31 structures, and ($\overline{SA}$)_r is the energy-minimized mean atomic structure.

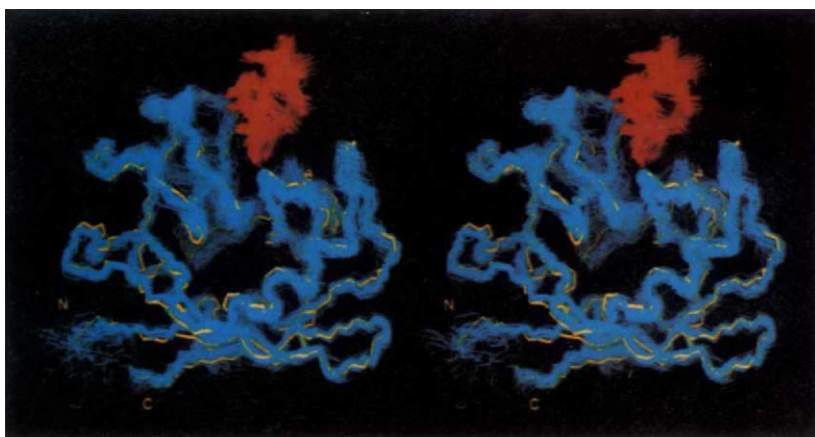
† None of the structures had distance violations ≥ 0.4 Å. Intraligand restraints included negative NOEs as previously described¹¹.

‡ XPLOR NOE energies were calculated using a square-well potential for the NOE empirical energy term with a force constant of $50 \text{ kcal mol}^{-1} \text{ Å}^{-2}$. The Lennard-Jones energy ($E_{\text{L-J}}$) was calculated using the CHARMM²⁷ empirical energy function.

§ R.m.s. deviation between the 31 structures and the mean atomic coordinates for CsA and CyP residues 5-165.

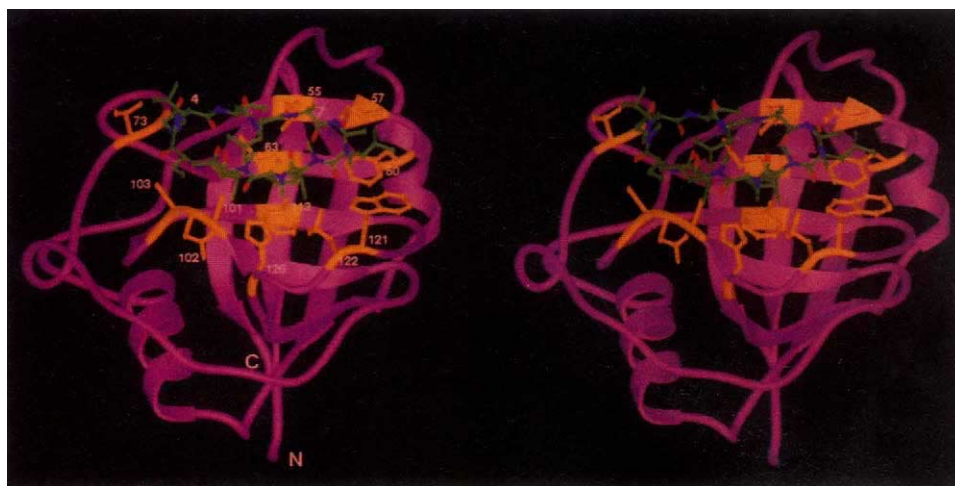
|| R.m.s. deviation between the 31 structures and the X-ray coordinates of uncomplexed CyP⁷ for CyP residues 5-165.

¶ R.m.s. deviation for residues CyP residues 5-165 between the minimized mean atomic solution structure and the X-ray coordinates of uncomplexed CyP⁷.



potential was used for the van der Waals term (F_{repel})¹⁷. Force constants for the NOE-derived distance restraints were maintained at $50 \text{ kcal mol}^{-1} \text{ Å}^{-2}$ throughout the refinement process. Dihedral angle restraints were initialized to $5 \text{ kcal mol}^{-1} \text{ rad}^{-2}$ and increased to $200 \text{ kcal mol}^{-1} \text{ rad}^{-2}$ at the beginning of annealing, which proceeded stepwise from 2,000 to 100 K in 50 K increments, with each step consisting of 0.25 ps of restrained molecular dynamics. The repulsive term was increased during this phase to $4.0 \text{ kcal mol}^{-1} \text{ Å}^{-4}$ and was maintained at that value during the penultimate energy minimization. Final structures were obtained after 1,000 cycles of Powell minimization with the full CHARMM Lennard-Jones potential²⁷.

FIG. 2 Ribbon plot³¹ depicting the structure of the CsA/CyP complex. Side-chain atoms are shown for those residues of cyclophilin (yellow) with NOEs to CsA colour-coded by atom type.



The structure of the CsA/CyP complex is stabilized by several hydrophobic interactions. The side chains of CyP residues Trp 121, Phe 60, Ile 57, Leu 122, Phe 113, His 126, Ala 101, Ala 103 and Thr 73 form a hydrophobic pocket that interacts with the hydrophobic surface of CsA residues 9–11 and 1–3 (Fig. 2). Several hydrogen bonds also stabilize the structure of the complex (Fig. 3a). The NMR data indicate intermolecular hydrogen bonds between MeLeu 10(CO) and Arg 55(η NH), MeLeu 9(CO) and Trp 121(ϵ NH), as well as Abu 2(NH) (where Abu represents L- α -aminobutyric acid) and Asn 102(CO), consistent with the slow exchange of the Abu 2 amide¹⁰.

Cyclosporin A binds to the same region of CyP as a tetrapeptide substrate (N-acetyl-Ala-Ala-Pro-Ala-amidomethylcoumarin), but in an opposite N- to C-terminal orientation (Fig. 2). As previously proposed¹³, the MeVal 11 side chain fits into the same hydrophobic pocket as the Pro ring of the substrate, and the MeVal 11 α -C' bond of CsA is superimposable on the N-C' bond of the substrate that isomerizes. Unlike the peptide bond of the tetrapeptide substrate, the MeVal 11 α -C' bond of CsA can more easily rotate to form favourable interactions with CyP, suggesting that this portion of CsA mimics a

twisted amide bond as postulated in one mechanism²⁴. These results contradict our earlier proposal²⁵ based on a model CsA/CyP structure and an unrefined X-ray structure⁸, in which the MeBmt 1OH mimics the carbonyl of the peptide bond (MeBmt is N-methyl-(4R)-4-((E)-2-butenyl)-4-methyl-L-threonine).

Four models of the CsA/CyP complex were reported based on X-ray and NMR data. The NMR structure of CsA/CyP was found to be very different from one of the proposed models of the complex¹², differing in the CsA orientation relative to CyP by about 90°. Three other models^{13,14,25} based on additional CsA/CyP NOE data are more similar to the NMR structure, but still differ in several important aspects. In contrast to two of the models^{13,14}, the proposed hydrogen bond between the MeBmt 1OH and Asn 102 carbonyl is not present in the NMR structure and cannot be used to explain the importance of the MeBmt 1OH group for cyclophilin binding and immunosuppressive activity¹⁶. Instead, the MeBmt 1OH hydrogen-bonds to the MeBmt 1 or MeLeu 4 carbonyl group and may be important for stabilizing the conformation of CsA. In addition, the Abu2 NH of CsA is hydrogen-bonded to Asn 102(CO) in our

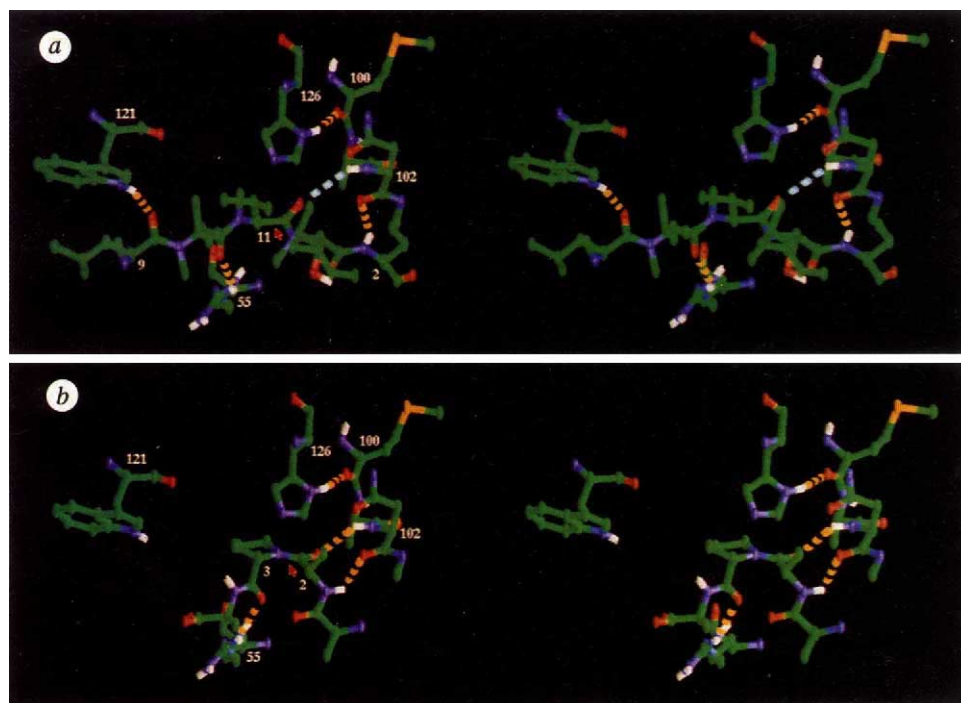


FIG. 3 Portions of *a*, the NMR structure of the CsA/CyP complex compared with *b*, a model of the X-ray structure of the tetrapeptide/CyP complex⁹, indicating similar hydrogen-bonding patterns (yellow dashed line). A direct or water-mediated hydrogen bond between Asn 102 NH and MeVal 11(CO) of CsA (blue dashed line) is also possible, as evidenced by the close proximity of these atoms in many of the NMR structures and the slow amide-exchange rate of Asn 102. This hydrogen bond would mimic the Asn 102 NH/Ala2(CO) hydrogen bond in the tetrapeptide/CyP complex shown in *b*. Arrows indicate the α -C' of CsA in *a* that corresponds to the N-C' bond of the tetrapeptide substrate in *b*.

NMR structure and not to MeBmt1 OH as previously reported^{13,14}, and appears to mimic the Ala 2(NH)/Asn 102(CO) hydrogen bond in the tetrapeptide/cyclophilin complex⁹ (Fig. 3). Furthermore, the proposed hydrogen bond between His 126 ϵ NH and MeVal 11(CO)^{13,14} is not possible in our NMR structure. Titrations show that at pH 6.5, His 126 is uncharged and only the δ (not ϵ) NH (12.51 p.p.m.), which is hydrogen-bonded to Met100(CO), is observed. Owing to the lack of NOEs involving Gln 63, a hydrogen bond between Gln ϵ NH and MeBmt 1(CO)¹³ could not be unambiguously identified.

The NMR structure is consistent with previously observed structure/activity relationships. For example, the marked decrease in CyP affinity for CsA analogues substituted with larger side chains at the 11 position^{15,16} can be rationalized by

the tight fit of the MeVal 11 in the middle of the binding pocket. In contrast, the side chains of MeLeu 9 and Abu 2 are located to the sides of the binding pocket (Fig. 2) and can accommodate a variety of side chains consistent with the structure/activity relationships^{15,16}. Cyclosporin A residues 5–8 do not interact with cyclophilin and, as expected from the NMR structure, are able to tolerate modifications. Some of the analogues with modifications at these sites bind tightly to cyclophilin but are relatively inactive as immunosuppressants (for example, [MeAla 6]CsA and [D-aminobutyric acid]CsA)^{8,15} which may be explained by their decreased affinity for calcineurin²⁶. Key CyP mutants together with the NMR structure of the CsA/CyP complex presented here should allow the identification of those portions of the complex that interact with calcineurin. □

Received 18 August; accepted 9 November 1992.

1. Takahashi, N., Hayano, T. & Suzuki, M. *Nature* **337**, 473–475 (1989).
2. Fischer, G., Wittmann-Liebold, B., Lang, K., Kiefhaber, T. & Schmid, F. X. *Nature* **337**, 476–478 (1989).
3. Handschumacher, R. E., Harding, M. W., Rice, J., Drugge, R. J. & Speicher, D. W. *Science* **226**, 544–547 (1984).
4. Liu, J. *et al.* *Cell* **66**, 807–815 (1991).
5. O'Keefe, S. J., Tamura, J., Kincaid, R. L., Tocci, M. J. & O'Neil, E. A. *Nature* **357**, 692–694 (1992).
6. Clipstone, N. A. & Crabtree, G. R. *Nature* **357**, 695–697 (1992).
7. Ke, H., Zydowsky, L. D., Liu, J. & Walsh, C. T. *Proc. natn. Acad. Sci. U.S.A.* **88**, 9483–9487 (1991).
8. Kallen, J. *et al.* *Nature* **353**, 276–279 (1991).
9. Kallen, J. & Walkinshaw, M. D. *FEBS Lett.* **300**, 286–290 (1992).
10. Weber, C. *et al.* *Biochemistry* **30**, 6563–6574 (1991).
11. Fesik, S. W. *et al.* *Biochemistry* **30**, 6574–6583 (1991).
12. Schreiber, S. L. & Crabtree, G. R. *Immun. Today* **13**, 136–142 (1992).
13. Spitzfaden, C. *et al.* *FEBS Lett.* **300**, 291–300 (1992).
14. Gallion, S. & Ringe, D. *Protein Engng* **5**, 391–397 (1992).
15. Sigal, N. H. *et al.* *J. exp. Med.* **173**, 619–628 (1991).
16. Quesniaux, V. F. J. *et al.* *Eur. J. Immun.* **17**, 1359–1365 (1987).
17. Nilges, M., Clore, G. M. & Gronenborn, A. M. *FEBS Lett.* **229**, 317–324 (1988).

18. Wüthrich, K., Spitzfaden, C., Memmert, K., Widmer, H. & Wider, G. *FEBS Lett.* **285**, 237–247 (1991).
19. James, M. N. G., Sielecki, A., Salituro, F., Rich, D. H. & Hofmann, T. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6137–6141 (1982).
20. Loosli, H. R. *et al.* *Helv. chim. Acta* **68**, 682–704 (1985).
21. Kessler, H., Köck, M., Wein, T. & Geherke, M. *Helv. chim. Acta* **72**, 1818–1832 (1990).
22. Altschuld, D., Vix, O., Rees, B. & Thierry, J.-C. *Science* **256**, 92–94 (1992).
23. Kofron, J. L. *et al.* *J. Am. chem. Soc.* **114**, 2670–2675 (1992).
24. Harrison, R. K. & Stein, R. L. *Biochemistry* **29**, 1684–1689 (1990).
25. Fesik, S. W., Neri, P., Meadows, R., Olejniczak, E. T. & Gemmecker, G. *J. Am. chem. Soc.* **114**, 3165–3166 (1992).
26. Liu, J. *et al.* *Biochemistry* **31**, 3896–3901 (1992).
27. Brooks, B. R. *et al.* *J. Comput. Chem.* **4**, 187–193 (1983).
28. Kuszewski, J., Nilges, M. & Brünger, A. T. *J. Biomolec. NMR* **2**, 33–56 (1992).
29. Neri, P. *et al.* *FEBS Lett.* **294**, 81–88 (1991).
30. Neri, D., Szyperski, T., Otting, G., Senn, H. & Wüthrich, K. *Biochemistry* **28**, 7510–7516 (1989).
31. Carson, M. *J. molec. Graphics* **5**, 103–106 (1987).

ACKNOWLEDGEMENTS. We thank R. Helfrich and J. Severin for preparing the CsA/[U-¹⁵N]CyP and CsA/[U-¹⁵N, ¹³C]CyP samples and H. Ke for atomic coordinates of uncomplexed cyclophilin. This work was partially supported by a grant from the NIH.

X-ray structure of a decameric cyclophilin–cyclosporin crystal complex

Gaston Pflügl*, Jörg Kallen, Tilman Schirmer*, Johan N. Jansonius*, Mauro G. M. Zurini & Malcolm D. Walkinshaw†

Preclinical Research, Sandoz Pharma AG, 4002-Basel, Switzerland
* Biocentre, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

HUMAN cyclophilin A (CypA), a ubiquitous intracellular protein of 165 amino acids, is the major receptor for the cyclic undecapeptide immunosuppressant drug cyclosporin A (CsA)^{1,2}, which prevents allograft rejection after transplant surgery^{3,4} and is efficacious in the field of autoimmune diseases⁵. CsA prevents T-cell proliferation by blocking the calcium-activated pathway leading to interleukin-2 transcription. Besides their ability to bind CsA, the cyclophilin isoforms^{6–8} also have peptidyl–prolyl isomerase activity^{9–11} and enhance the rate of protein folding^{12,13}. The macrolide FK506 acts similarly to CsA and its cognate receptor FKBP also has peptidyl–prolyl isomerase activity¹⁴. Inhibition of this enzymatic activity alone is not sufficient to achieve immunosuppression^{15,16}. A direct molecular interaction between the drug–immunophilin complex (CsA–CypA, or FK506–FKBP) and the phosphatase calcineurin, is responsible for modulating the T-cell receptor signal transduction pathway^{17,18}. Here we describe the crystal structure of a decameric CypA–CsA complex. The crystallographic asymmetric unit is composed of a pentamer of 1:1 cyclophilin–cyclosporin complexes of rather exact non-crystallographic fivefold symmetry. The 2.8 Å electron density map is of

high quality. The five independent cyclosporin molecules are clearly identifiable, providing an unambiguous picture of the detailed interactions between a peptide drug and its receptor. It broadly confirms the results of previous NMR, X-ray and modelling studies, but provides further important structural details which will be of use in the design of drugs that are analogues of CsA.

The crystal structure of CypA has been solved as a complex with a peptidyl–prolyl isomerase (PPIase)-substrate analogue tetrapeptide^{19,20} and in an uncomplexed form²¹. NMR studies of the CypA–CsA complex revealed the bound conformation of CsA^{22,23}. Subsequent modelling studies provided a preliminary structure of CsA docked into CypA^{24–26}. A number of other crystal forms of the CypA–CsA complex have large unit cells with multiple copies of the complex in the asymmetric unit²⁷. We have solved the tetragonal crystal form with a pentamer of the complex in the asymmetric unit (Fig. 1a; see legend for structure determination). The fivefold rotational symmetry of the pentamer is quite exact and even some 250 of the 509 positioned water molecules obey the fivefold symmetry. The average root-mean-square deviation (r.m.s.d.) between the five CypA molecules is 0.59 Å (for all pairwise fits using all atoms). The refined X-ray structure of CypA bound to a tetrapeptide which was used in molecular replacement has an r.m.s.d. of 1.1 Å (all atoms) from the pentamer protomers. The five independent CsA molecules are also closely similar and have an average r.m.s.d. of 0.26 Å (for all pairwise fits using all atoms).

The cyclophilin pentamers pack in pairs to form decamer discs which sandwich the CsA molecules. A side view of the decamer (Fig. 1b) shows that CsA-binding loops of CypA are on one face of the pentamer whereas the secondary structure elements line up on the opposite surface. In the crystal form described here, a (crystallographic) twofold rotation axis intersects the fivefold axis orthogonally (Fig. 2). The environment of each CypA molecule in the 'decameric cyclophilin–cyclosporin sandwich' is very similar, with protein–protein interactions made to four CypA neighbours (Fig. 2).

† To whom correspondence should be addressed.

The CsA backbone conformation is similar to that determined from NMR studies of the complex²², with an r.m.s.d. of 0.55 Å (for C α atoms of CsA) between the crystal complex and the modelled docked structure²⁴. The major new finding is the existence of an intramolecular hydrogen bond between the hydroxyl group of MeBmt-1 and the carbonyl oxygen of MeLeu-4. The docked structure had this hydroxyl group pointing towards CypA forming a hydrogen bond with the carbonyl oxygen of Asn 102 (ref. 24).

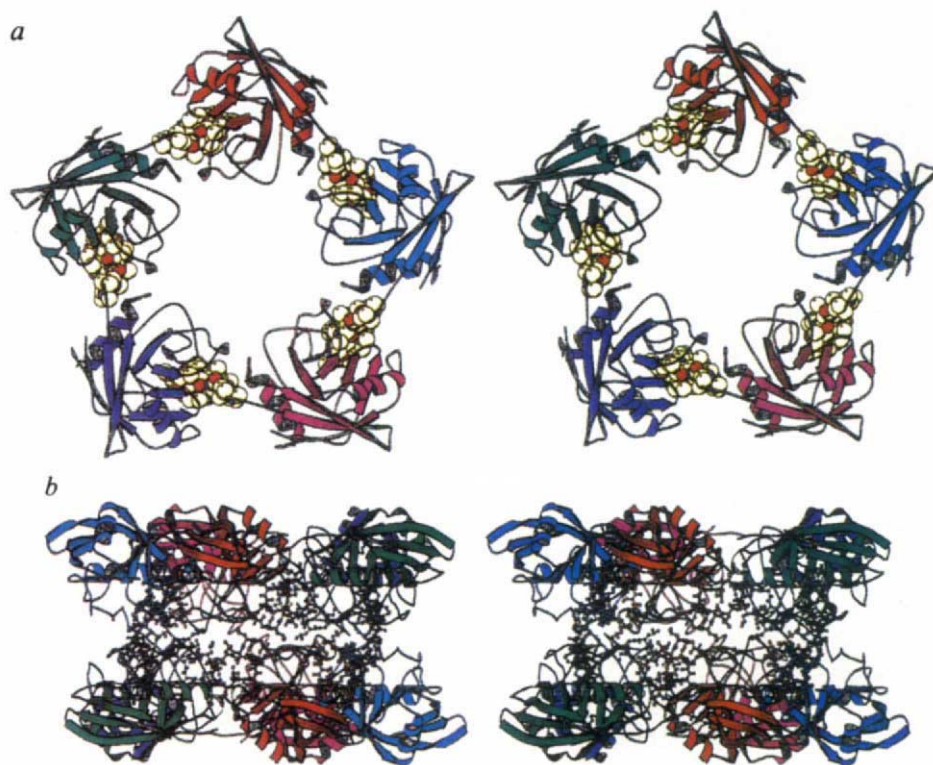
The crystal structure shows that CsA binds to CypA with the side chains of MeBmt-1, Abu-2, Sar-3, MeLeu-9, MeLeu-10 and MeVal-11 of CsA making contacts of less than 4 Å with the CypA residues Arg 55, Phe 60, Met 61, Gln 63, Gly 72, Ala 101, Asn 102, Ala 103, Gln 111, Phe 113, Trp 121, Leu 122, His 126 (Fig. 3a). These thirteen CypA residues define the CsA binding site and are identical to those proposed by the modelled docked structure²⁴. Nine of these residues are also within 4 Å of the tetrapeptide acetyl-Ala-Ala-Pro-Ala-amc (amc is amidomethylcoumarin) bound in the active site²⁰. The active-site residues of

the two complexes have very similar conformations. The largest side-chain movements are 1.3 Å for Arg 55 and up to 0.7 Å for Phe 60, Gln 63 and Trp 121. The fit between CypA and the binding edge of CsA is excellent and centres round the protruding side chain of MeVal-11. The solvent-accessible surface area of free CsA in this bound conformation is 1,238 Å², of which 42% (526 Å²) becomes inaccessible in the CypA-CsA protomer, as measured with the program AREAIMOL (SERC, Daresbury, UK). There are four direct hydrogen bonds between CsA and CypA (Table 1; Fig. 3a). In the modelled docked structure²⁴, residues 4, 5, 6, 7, 8 of CsA protrude out into the solvent, and are thought to be involved in binding the effector protein, calcineurin. In the decamer, this so-called effector region of CsA interacts with two other CypA molecules as well as a (twofold related) CsA molecule (Fig. 3b). There are two direct hydrogen-bonds from neighbouring CypA molecules in the 'decameric sandwich' to the effector region of each CsA (Table 1; Figs 2 and 3b).

The formation of a supramolecular complex in the crystal

FIG. 1 a, Stereo drawing using MOLSCRIPT³² of the crystallographic asymmetric unit which consists of a planar pentamer complex of five CypA molecules bound to five CsA molecules. The red CypA molecule at the top of the diagram corresponds to molecule A in Fig. 2. CsA molecules are drawn in space-filling mode with carbons yellow, oxygens red and nitrogens blue. b, Stereo ribbon drawing³² of the 'decameric cyclophilin-cyclosporin sandwich' viewed along the crystallographic dyad and orthogonal to the (vertical) fivefold axis. The five cyclophilin pairs related by the crystallographic dyad are coloured identically. The 10 CsA molecules are in the middle of the sandwich interacting with CsA binding loops from the cyclophilin molecules. The CsA carbon atoms of each CsA are coloured the same as its bound CypA molecule.

METHODS. Details of X-ray structure determination: intensity data to 2.8 Å resolution were collected on tetragonal crystals (space group $P4_32_12$ with $a = b = 95.2$ Å, $c = 280.0$ Å) at the DESY Synchrotron Source (Hamburg) using an image plate detector (200,668 measurements of 32,027 unique reflections; R sym = 8.1%; completeness, 98%). Subsequently an additional data set to 5.0 Å was collected with a FR571 rotating anode using a FAST area detector (46,052 measurements of 11,333 unique reflections; $R_{\text{sym}} = 9.3\%$; completeness, 80.9%). Self-rotation function calculations with the program POLARRFN (SERC Daresbury Laboratory, UK) using data between 8 Å and 5 Å (Patterson radius 30 Å) yielded unique peaks at $\omega = 18$, $\phi = 45$, $\kappa = 71$ (3.8σ), at $\omega = 75$, $\phi = 26$, $\kappa = 180$ (3.5σ) and at $\omega = 80$, $\phi = 10$, $\kappa = 180$ (4.2σ). These peaks, together with a crystallographic twofold axis, were interpreted as possibly resulting from 52 point-group symmetry. Cross-rotation function calculations with XPLOR³³ using the refined model (without CsA) of CypA (ref. 20) and data between 10 Å and 3 Å gave 5 unique peaks (3.3 – 3.5σ) at $\omega = 29.7$, $\phi = 193.1$, $\kappa = 204.4$, at $\omega = 38.6$, $\phi = 184.2$, $\kappa = 249.1$, at $\omega = 26.4$, $\phi = 186.8$, $\kappa = 230.8$, at $\omega = 38.0$, $\phi = 179.3$, $\kappa = 188.6$, and at $\omega = 51.1$, $\phi = 201.6$, $\kappa = 250.3$. These peaks were consistent with the existence of a local fivefold rotation axis and also provided a calculated self-rotation function which agreed very well with the observed self-rotation function. Translation function calculations for each of the five cross-rotation function solutions using the program TFSGEN (SERC, Daresbury, UK) and data between 48 Å and 3.1 Å resulted in strong peaks (14.1 – 17.9σ) and indicated space group $P4_32_12$ rather than $P4_12_12$. The program TFPART (SERC, Daresbury, UK) was used to correctly position these five translation function solutions relative to each other (yielding subsequent peak heights increasing from



22.3 σ to 50.0 σ). The initial R -factor (without CsA) for data between 48 Å and 3.1 Å was 37.6% and a $F_o - F_c$ map showed excellent density for the five CsA molecules. Refinement with XPLOR using data between 48 Å and 2.8 Å gave an R -factor of 28.4%. CsA was then built into a $F_o - F_c$ map and subsequent refinement resulted in an R -factor of 21.3%. Individual B -factors were then refined and 509 water molecules added to give a final R -factor of 15.7% for all data between 20 Å and 2.8 Å (32,020 reflections). All refinement steps were done without imposing non-crystallographic symmetry constraints. For the final model the r.m.s. deviations from target values are 0.011 Å for bond lengths and 2.7° for bond angles. An estimate of how well the non-crystallographic fivefold symmetry is conserved has been made by applying an exact fivefold rotation for a given atom and determining the r.m.s.d. from the observed positions. For all CypA atoms, including side chains, the average r.m.s.d. between the observed and five-fold calculated positions, is 0.61 Å and for C α atoms this r.m.s.d. is 0.26 Å. The non-crystallographic fivefold symmetry is also strongly conserved between the five CsA molecules with average r.m.s.d. values of 0.27 Å for all atoms and 0.15 Å for the eleven C α atoms. The conservation of the fivefold symmetry also holds for 250 of the water molecules which show an average r.m.s.d. from exact fivefold symmetry of 0.3 Å.

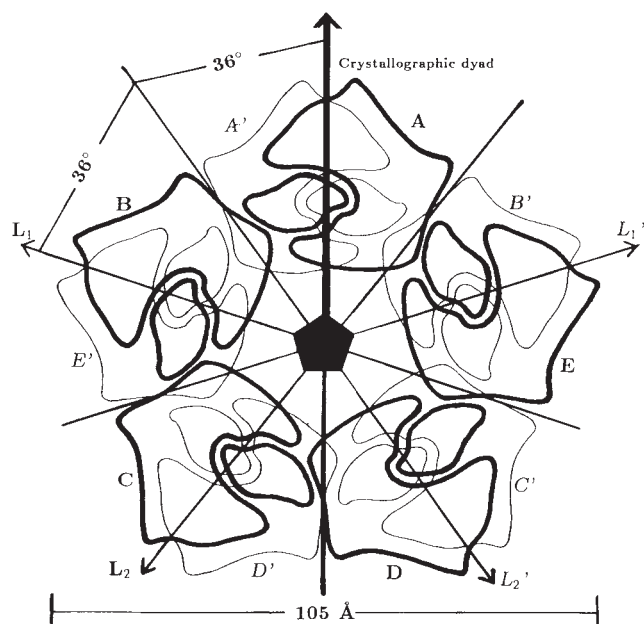


FIG. 2 Diagrammatic representation of the 'decameric cyclophilin-cyclosporin sandwich' viewed along the fivefold axis. The top face of the decamer is composed of the pentamer labelled A, B, C, D, E and the bottom symmetry-related face is labelled A', B', C', D', E'. The decamer has a diameter of ~ 105 Å and a depth of ~ 66 Å. The hole through the middle has a diameter of ~ 16 Å. The crystallographic dyad $[1\ 1\ 0]$ orthogonally intersects the fivefold axis which lies in the (110) plane and is inclined 16.8° from the c axis. The two unique noncrystallographic twofold axes (L_1 and L_2 which were major peaks in the self-rotation function) are shown. The diagram also indicates the intermolecular contacts: A is in contact with E, B, A' and B', while CsA-A makes contacts with A, B, A' and CsA-A'.

FIG. 3 *a*, Stereo drawing³² of one of the five CsA-CypA complexes. The ribbon representation of CypA has helices coloured green, strands coloured red and turns coloured pale blue. Those 13 CypA residues involved in CsA binding are also drawn (see text). CsA has the sequence c-(MeBmt-Abu-Sar-MeLeu-Val-MeLeu-Ala-D-Ala-MeLeu-MeLeu-MeVal), where Me indicates *N*-methylation and MeBmt is (4*R*)-4-[(*E*)-2-butenyl]-4-*N*-dimethylthreonine. CsA is coloured with carbons grey, oxygens red, nitrogens blue and all bonds yellow. CsA residue numbers are shown. The four CsA...CypA hydrogen bonds are drawn as dashed black lines (see Table 1). The intramolecular hydrogen bond between the hydroxyl group of MeBmt and the carbonyl oxygen of MeLeu4 is also shown. *b*, Stereo drawing³² of the environment of a CsA molecule in the 'decameric sandwich'. The central CsA molecule is drawn with bonds coloured yellow. CypA residues which have any atom < 4 Å from any atom of the central CsA molecule are drawn. Hydrogen bonds are shown as dashed lines. No water molecules are included. Residues 9, 10, 11, 1, 2 of CsA sit in the binding pocket of the blue CypA molecule. Residues 2, 3, 4 are within 4 Å of a dyad-related CsA molecule drawn here with green carbons and pink bonds. CsA residues 4, 6, 7, 8 and 9 in the effector region make contact to two CypA molecules in the decameric sandwich (coloured here in green and red). The two direct hydrogen bonds in the effector region from Ala 7 to Lys 31 and Glu 81 are shown.

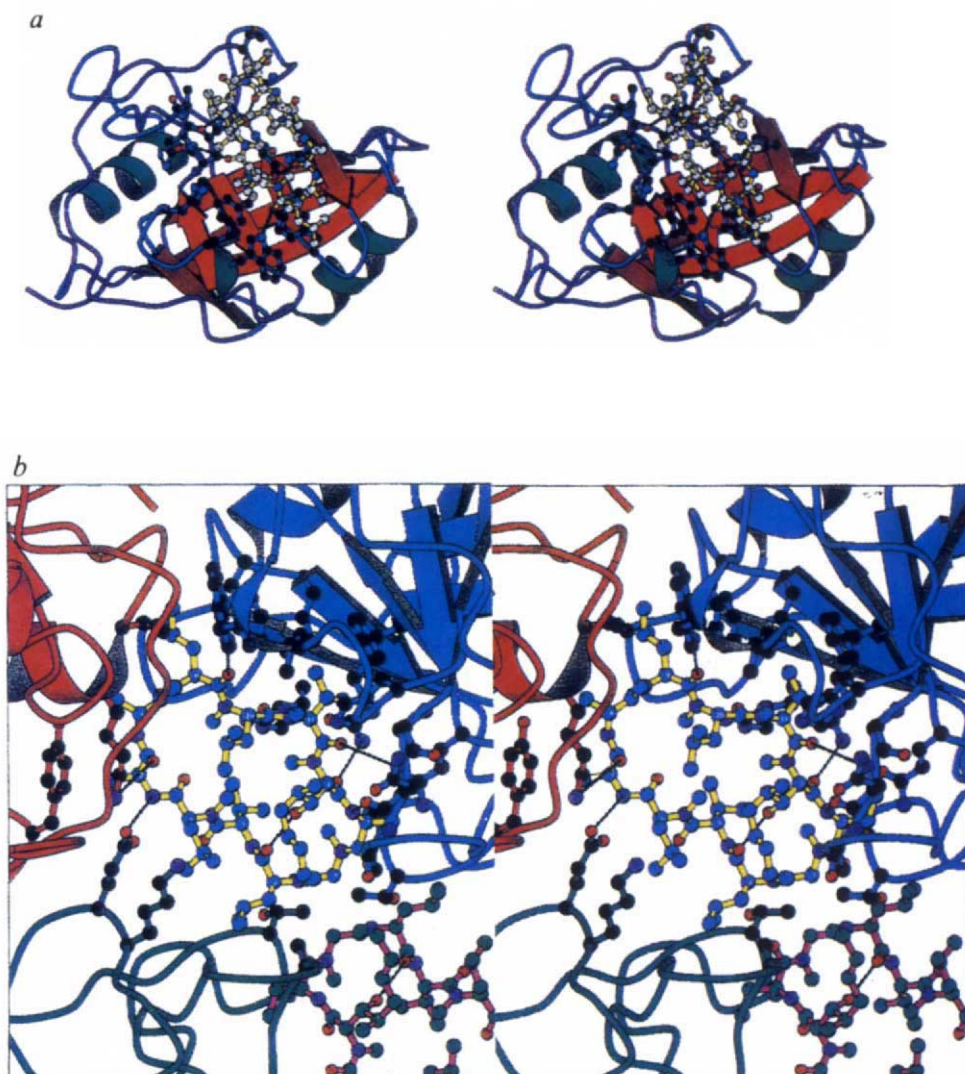


TABLE 1 Intermolecular hydrogen bonds within the decameric cyclophilin-cyclosporin sandwich

From CypA molecule A to molecules E and B within the pentamer		
CypA atom	CypA atom	Å(range)
A28 Lys NZ ... E120 Glu OE		3.05 (2.43–3.10)
A43 Glu OE ... E148 Arg NH		3.90 (3.33–3.90)
A120 Glu OE ... B28 Lys NZ		2.50 (2.43–3.10)
A148 Arg NH ... B43 Glu OE		3.90 (3.33–3.90)
From CypA molecule A to molecule B'		
A31 Lys NZ ... B'81 Glu OE		3.29 (2.60–3.29)
A79 Tyr O ... B'81 Glu N		3.18 (3.00–3.18)
A81 Glu N ... B'79 Tyr O		3.00 (3.00–3.18)
A82 Lys NZ ... B'84 Glu OE		3.49 (3.32–3.99)
A84 Glu OE ... B'82 Lys NZ		3.99 (3.32–3.99)
Hydrogen bonds between CsA and CypA		
CsA atom	CypA atom	Å(range)
A1 MeBmt O ... A63 Gln NE		2.94 (2.94–3.27)
A2 Abu N ... A102 Asn O		3.25 (2.92–3.25)
A9 MeLeu O ... A121 Trp NE		2.81 (2.81–2.96)
A10 MeLeu O ... A55 Arg NH		3.06 (2.77–3.06)
A11 MeVal O ... A126 His CE		3.40 (3.28–3.67)*
A7 Ala O ... B31 Lys NZ		3.93 (2.84–3.93)
A7 Ala N ... A'81 Glu OE		3.11 (2.80–3.11)
Hydrogen bonds between CsA and water (W)		
CsA atom		Å (range)
A1 MeBmt O ... W1		3.1 (3.1–3.2)
A5 Val O ... W5		3.0 (3.0–3.5)
A6 MeLeu O ... W6		3.4 (2.3–3.6)
A8 D-Ala N ... W8		3.3 (2.9–3.3)

Each of the five independent CypA molecules (A, B, C, D, E) in the planar pentamer form very similar contacts. The atom name consists of the molecule name (Fig. 2), residue number, residue type and atom type. The interatomic distance (Å) is given for one of the five contacts. The range of distances found for the other four molecules is given in parentheses.

* The carbonyl oxygen atom of MeVal-11 of CsA is about 3.4 Å from CE of His 126. Hydrogen-bond formation between CsA and His 126 would require 180° rotation about χ_2 , which would however preclude the formation of two alternative His 126 hydrogen bonds.

raises the question of the possible biological relevance of such oligomers. Unit cell volumes of two other crystal forms of CypA-CsA complexes are consistent with ten molecules per asymmetric unit²⁷. This suggests that the 'decameric sandwich' may exist in solution. An immunophilin of *M*, 59K is known to associate with two heat-shock proteins (Hsp 90 and Hsp 70) and the glucocorticoid receptor²⁸. Heat-shock proteins mediate protein folding²⁹ and some function as oligomers. For example, the Hsp 60 chaperonins are composed of two stacked 7-subunit rings. Two other protein crystal structures, the 23K human C-reactive protein³⁰ and a homologous serum amyloid P component³¹, were found to be pentameric. In both cases this may be biologically relevant. In the 'decameric sandwich', CsA has about 80% of its surface hidden, leaving a solvent-accessible surface of 270 Å²; the CsA binding site (and therefore the PPIase active site) is completely inaccessible. □

Received 17 August; accepted 12 November 1992.

- Harding, M. W., Handschumacher, R. E. & Speicher, D. W. *J. biol. Chem.* **261**, 8547–8552 (1986).
- Handschumacher, R. E., Harding, M. W., Rice, J., Drugge, R. J. & Speicher, D. W. *Science* **226**, 544–547 (1984).
- Calne, R. Y. *et al. Lancet* **2**, 1323–1326 (1978).
- Borel, J. F. *Pharmac. Rev.* **41**, 259–371 (1989).
- Feutren, G. *Transplant. Proc.* **24** (suppl. 2), 55–60 (1992).
- Swanson, S. K.-H. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 3741–3745 (1992).
- Price, E. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 1903–1907 (1991).
- Friedman, J. & Weissman, I. *Cell* **66**, 799–806 (1991).
- Harrison, R. K. & Stein, R. L. *J. Am. chem. Soc.* **114**, 3464–3471 (1992).
- Kofron, J. L., Kuzmic, P., Kishore, V., Colon-Bonilla, E. & Rich, D. H. *Biochemistry* **30**, 6127–6134 (1991).

- Fischer, G., Wittmann-Liebold, B., Lang, K., Kiefhaber, T. & Schmid, F. X. *Nature* **337**, 476–478 (1989).
- Schoenbrunner, E. R. *et al. J. biol. Chem.* **266**, 3630–3635 (1991).
- Fransson, C. *et al. FEBS Lett.* **296**, 90–94 (1992).
- Rosen, M. K. & Schreiber, S. L. *Angew. Chem.* **104**, 413–430 (1992).
- Schreiber, S. L. *Science* **251**, 283–287 (1991).
- Sigal, N. H. & Ali, E. *J. exp. Med.* **173**, 619–628 (1991).
- O'Keefe, S. J., Tamura, J., Kincaid, R. L., Tocci, M. J. & O'Neill, E. A. *Nature* **357**, 692–694 (1992).
- Clipstone, N. A. & Crabtree, G. R. *Nature* **357**, 695–697 (1992).
- Kallen, J. *et al. Nature* **353**, 276–279 (1991).
- Kallen, J. & Walkinshaw, M. D. *FEBS Lett.* **300**, 286–290 (1992).
- Ke, H., Zydowsky, L. D., Liu, J. & Walsh, C. T. *Proc. natn. Acad. Sci. U.S.A.* **88**, 9483–9487 (1991).
- Weber, C. *et al. Biochemistry* **30**, 6563–6574 (1991).
- Fesik, S. W. *et al. Biochemistry* **30**, 6574–6583 (1991).
- Spitzfaden, C. *et al. FEBS Lett.* **300**, 291–300 (1992).
- Fesik, S. W., Neri, P., Meadows, R., Olejniczak, E. T. & Gemmecker, G. *J. Am. chem. Soc.* **114**, 3165–3166 (1992).
- Gallion, S. & Ringe, D. *Protein. Engng* **5**, 391–397 (1992).
- Zurini, M. *et al. FEBS Lett.* **276**, 63–66 (1990).
- Tai, P. K., Albers, M. W., Chang, H., Faber, L. E. & Schreiber, S. L. *Science* **256**, 1315–1318 (1992).
- Langer, T. *et al. Nature* **356**, 683–689 (1992).
- Myles, D. A. A. *et al. J. molec. Biol.* **216**, 491–496 (1990).
- Wood, S. P. *et al. J. molec. Biol.* **202**, 169–173 (1988).
- Lindqvist, Y. *J. Molec. Biol.* **209**, 151–166 (1989).
- Brunger, A. T., Kuriyan, J. & Karplus, M. *Science* **235**, 458–460 (1987).

Active-centre torsion-angle strain revealed in 1.6 Å-resolution structure of histidine-containing phosphocarrier protein

Zongchao Jia, Margaret Vandonselaar*, J. Wilson Quail & Louis T. J. Delbaere*

Department of Chemistry, University of Saskatchewan, Saskatoon, Saskatchewan, S7N 0W0 Canada

* Department of Biochemistry, University of Saskatchewan, Saskatoon, Saskatchewan, S7N 0W0 Canada

THE histidine-containing phosphocarrier protein (HPr) is a central component of the phosphoenolpyruvate: sugar phosphotransferase system that transports carbohydrates across the cell membrane of bacteria¹. A typical phosphotransfer sequence is phosphoenolpyruvate → enzyme I → HPr → enzyme II/III^{sugar} → sugar. This is thermodynamically favourable owing to the participation of the high-energy phosphoenolpyruvate. We report here the structure of HPr from *Streptococcus faecalis* determined at 1.6 Å resolution. Remarkable disallowed Ramachandran torsion angles at the active centre, revealed by the X-ray structure, demonstrate a unique example of torsion-angle strain that is probably directly involved in protein function. During phosphorylation, the active-centre torsion-angle strain should facilitate the phosphotransfer reaction by lowering the activation-energy barrier. A recently reported *Bacillus subtilis* HPr structure², which represents the phosphorylated state of HPr with no torsion-angle strain, provides direct evidence supporting our hypothesis that torsion-angle strain plays a direct part in the function of HPr. An HPr phosphotransfer cycling mechanism is proposed, based primarily on the structures of HPr and other phosphotransferase system proteins.

The structure of HPr from *S. faecalis* has been solved using multiple isomorphous replacement (MIR) followed by solvent flattening³ and has been least-squares refined at 1.6 Å resolution. Details of crystallization, data collection, MIR phase determination, structure refinement and overall structure description are given in Table 1 and the legend to Fig. 1. The active centre of HPr consists of a histidine residue, at position 15 (His 15), the target of phosphorylation, and Arg 17. In *S. faecalis* HPr, the side chains of His 15 and Arg 17 point away from each other to adopt an open conformation (Fig. 2a) in the phospho-free state. This conformation is rigidly maintained by three hydrogen bonds (see legend to Fig. 2a). His 15 in *S. faecalis* HPr has a relatively low pK_a (6.1)⁴ and it was generally accepted that this resulted from the direct interaction between the side chains of

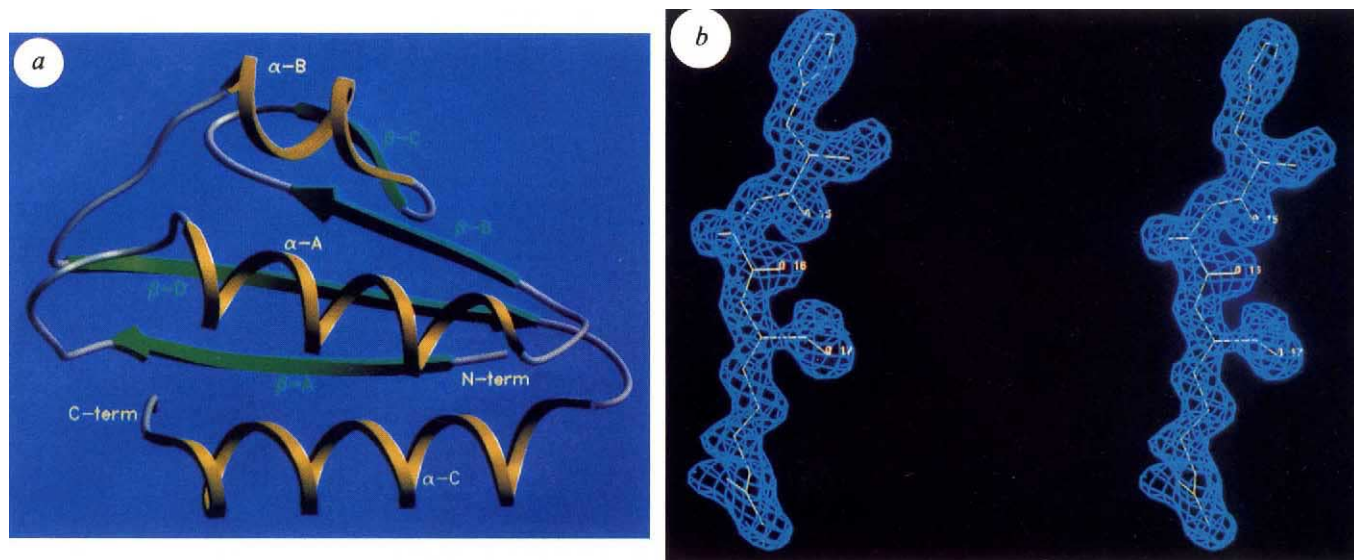


FIG. 1 *a*, Schematic ribbon representation of the overall folding of *S. faecalis* HPr; the ribbon illustration was generated using the program SETOR¹⁸. The overall folding topology of HPr is a classical open-faced β -sandwich. The β -sheet with right-handed twist consists of four antiparallel β -strands and the helical face has two major antiparallel α -helices and a minor α helix. The three α -helices are α -A(16–27), α -B(46–52) and α -C(69–84); the four β -strands are β -A(2–8), β -B(32–36), β -C(40–44) and β -D(58–66). A comparison of the NMR structures of HPr from *B. subtilis*¹⁹ and HPr from *E. coli*^{20,21} with this structure shows that they are very similar in overall tertiary folding, although the minor α -helix was not observed in the NMR structures. The X-ray structure of *S. faecalis* HPr is in general agreement with all HPr structures published so far^{2,19–23}, except for the *E. coli* HPr X-ray structure²⁴. Therefore the general HPr structural motif in different bacteria is apparently the open-faced β -sandwich. There are two C-terminal

residues (Glu 88 and Gln 89) missing from the present mode. *b*, An omit ($2F_o - F_c$) stereo map of the active centre with the final model (His 15, Ala 16 and Arg 17) superimposed; for the omit map calculation, residues 15–17 were left out of both the least-squares refinement and the map calculation; the map is contoured at a 1.0σ level.

METHODS. Residues 1–78 of HPr were fitted in a 2.8 Å solvent-flattened MIR map using TURBO-FRODO (ref. 9; Biographics, Marseilles). The partial model was refined using X-PLOR²⁵; residues 79–87 were then constructed on ($3F_o - 2F_c$) and ($F_o - F_c$) maps. The refinement was gradually extended to higher resolution. Manual adjustment of the model was carried out at intervals of the refinement. The *R* factor is 0.168 for all data between 8.0 and 1.6 Å resolution, maintaining good geometry with r.m.s. deviations in bond lengths of 0.012 Å and in bond angles of 2.22°. Coordinates will be deposited in the Brookhaven Protein Data Bank.

His 15 and Arg 17. Based on the current structure, however, this low pK_a can be attributed instead to the positive dipole moment at the N terminus of the α -A helix^{5,6} as His 15 N-caps the α -A helix and the side chains of His 15 and Arg 17 are far apart.

Evidence that this open conformation of the active centre is involved in protein function is provided by the site-directed mutagenesis study of HPr from *Escherichia coli*, in which it was found that a Gln 51 → Glu mutation decreased phosphotransferase activity by 90% (ref. 7). Modelling indicates that in the open conformation of *S. faecalis* HPr, if Met 51 were to be replaced by Gln 51, then the resulting longer side chain would be close enough to form a direct hydrogen bond to Nε2 of His 15. The Gln 51 → Glu mutation in *E. coli* HPr introduces an extra negative charge and would affect the electrostatic and/or conformational microenvironment of the active centre, resulting in a significant change in the biological activity of HPr. We propose that in the phospho-free state of HPr, the open conformation of the active centre (having residue 51 in the vicinity) is the one most likely to be involved in the phosphotransfer process with enzyme I. But the actual interaction awaits the structure determination of enzyme I and, more relevantly, structure determination of the enzyme I/HPr complex.

Only one residue in the *S. faecalis* HPr structure, Ala 16 (located in the middle of the active centre), adopts energetically unfavourable or disallowed Ramachandran⁸ angles (Fig. 2c). The strained conformation is illustrated in Fig. 1b, together with a section of the electron density map for which the tripeptide His 15-Ala 16-Arg 17 was omitted from all calculations. In addition, the ω angle of Arg 17 is 168°, 12° from planarity. We realized that these strained angles could be important in the phosphotransfer function of HPr. Modelling studies show that by rotation of main-chain torsion angles to achieve only allowed Ramachandran values (Fig. 2c) and adjustments of the side-

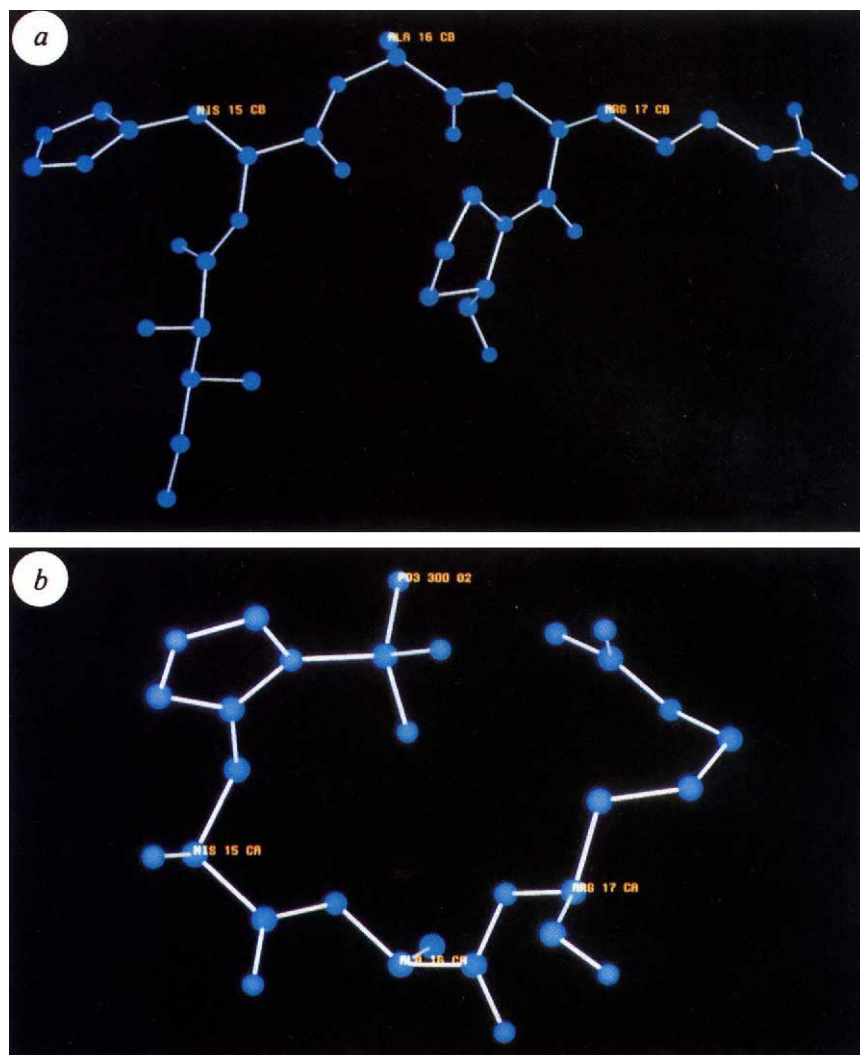
TABLE 1 Crystallographic and multiple isomorphous replacement data

Crystal	Native	KAuCl ₄	cis-Pt(NH ₃) ₂ Cl ₂
No. of measurements (number of crystals)	88,281 (2)	4,036 (1)	6,421 (1)
No. of unique reflections (completeness, %)	9,587 (95.0)	1,752 (78.5)	1,876 (84.9)
<i>R</i> _{sym}	0.061	0.062	0.037
Resolution limit (Å)	1.6	2.7	2.7
<i>R</i> _{iso}		0.391	0.090
Phasing power (<i>F</i> _H / <i>E</i>)		4.75	1.75
<i>R</i> _{cutlis}		0.360	0.651
Number of major sites		2	1
Mean figure of merit ($\langle m \rangle$)		0.86 at 2.8 Å resolution	

$R_{\text{sym}} = \sum |I - \langle I \rangle| / \sum I$; $R_{\text{iso}} = \sum ||F_{\text{PH}}| - |F_{\text{P}}|| / \sum |F_{\text{P}}|$; phasing power (F_{H}/E) = $\sqrt{\sum f_{\text{H}}^2 / \sum (F_{\text{PH}} - F_{\text{PH(calc)}})^2}$; $R_{\text{cutlis}} = \sum ||F_{\text{PH}} \pm F_{\text{P}}| - F_{\text{H(calc)}}| / \sum |F_{\text{PH}} - F_{\text{P}}|$ for centric reflections. *F*_P, *F*_{PH} and *F*_H, native, derivative and heavy-atom structure factors, respectively.

The *S. faecalis* HPr (*M*_r = 9,438) native crystals were grown by vapour diffusion from a solution containing 2.7 mg ml⁻¹ protein in citrate-phosphate buffer (pH 5.0) equilibrated against 45% saturated ammonium sulphate. Space group is *P*2₁2₁2, with cell dimensions *a* = 53.56, *b* = 45.46, *c* = 29.88 Å. The X-ray source was an Enraf-Nonius FR571 rotating-anode generator operated at 45 kV, 95 mA, with copper radiation and a graphite monochromator; data were collected on an Enraf-Nonius FAST area detector diffractometer. The temperature of measurement was 14 °C. Data were processed using the MADNES¹⁶ software package and the CCP4 suite (Daresbury Laboratory, UK) was used for all subsequent crystallographic calculations. Final heavy-atom parameters used for the phasing were obtained by a Dickerson-type refinement¹⁷, as implemented in the CCP4 program PHARE (G. Bricogne). Anomalous scattering data were included in the phasing calculation and were also used to determine the correct enantiomorph. A 2.8-Å MIR map was considerably improved using solvent flattening³, even though there is a relatively low solvent content (36%).

FIG. 2 *a*, Open, or phospho-free, form of the active centre in HPr. For clarity, the three hydrogen bonds mentioned in the text are not illustrated. CB represents $C\beta$ of an amino acid. The main chain O of Thr 12 forms a hydrogen bond with $N\delta 1$ of His 15 while $N\epsilon 2$ of His 15 hydrogen-bonds to a well-ordered water molecule that in turn hydrogen-bonds to the main chain O of Met 51. There is one water molecule hydrogen-bonded to $N\eta$ of Arg 17. *b*, Proposed model of closed, or phosphorylated, form of the active centre in HPr. CA represents $C\alpha$ of an amino acid. *c*, Ramachandran plot for the structure of *S. faecalis* HPr; crosses and squares represent non-glycine and glycine residues, respectively. The open-form torsion angles are observed values for the phospho-free state of HPr; the closed-form torsion angles of Ala 16 are for the modelled phosphorylated state of HPr.



chain torsion angles, the side chains of His 15 and Arg 17 can be easily brought together to make a closed conformation (Fig. 2*b*) suitable for accommodating a phosphoryl group; adjustments of main-chain and side-chain torsion angles were made using the program TURBO-FRODO (ref. 9; Biographics, Marseilles). Phospho-HPr (P-HPr) in the closed conformation is stabilized both by the positive dipole moment of helix A and by the positive charge of Arg 17. The energy released by going to allowed Ramachandran angles was estimated to be ~ 4 kcal $\cdot$ mol $^{-1}$ using the program BIOGRAF (BioDesign, Pasadena). This torsion-angle strain would lower the activation-energy barrier for the phosphorylation of His 15. Therefore, we suggest that this strain facilitates the phosphotransfer process from enzyme I to HPr. It should be noted that the local 'high energy' torsion-angle strain is compensated by favourable interactions such as the three hydrogen bonds, bringing the protein to its overall energy minimum. The energy of this hydrogen bonding¹⁰ is comparable to that of the torsion-angle strain. The breaking of hydrogen bonds in the active centre has been detected during phosphorylation of His 15 (ref. 11). Disallowed Ramachandran angles associated with active sites have been found in some proteins and are generally believed to have structural and also possible functional roles¹². To our knowledge, there are no other examples in which the activity/angle-strain relationship has been fully rationalized. Hence we present here the first rationalization of active-centre disallowed Ramachandran angles that are likely to play a direct role in protein function. The recently

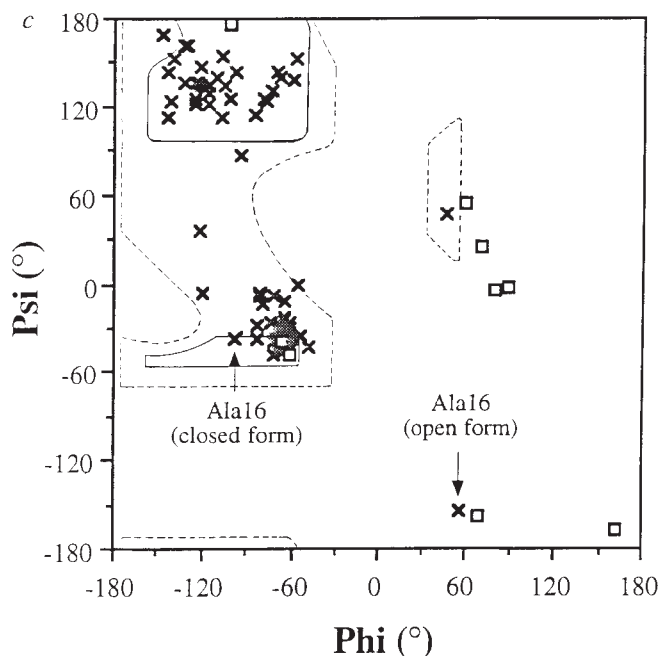
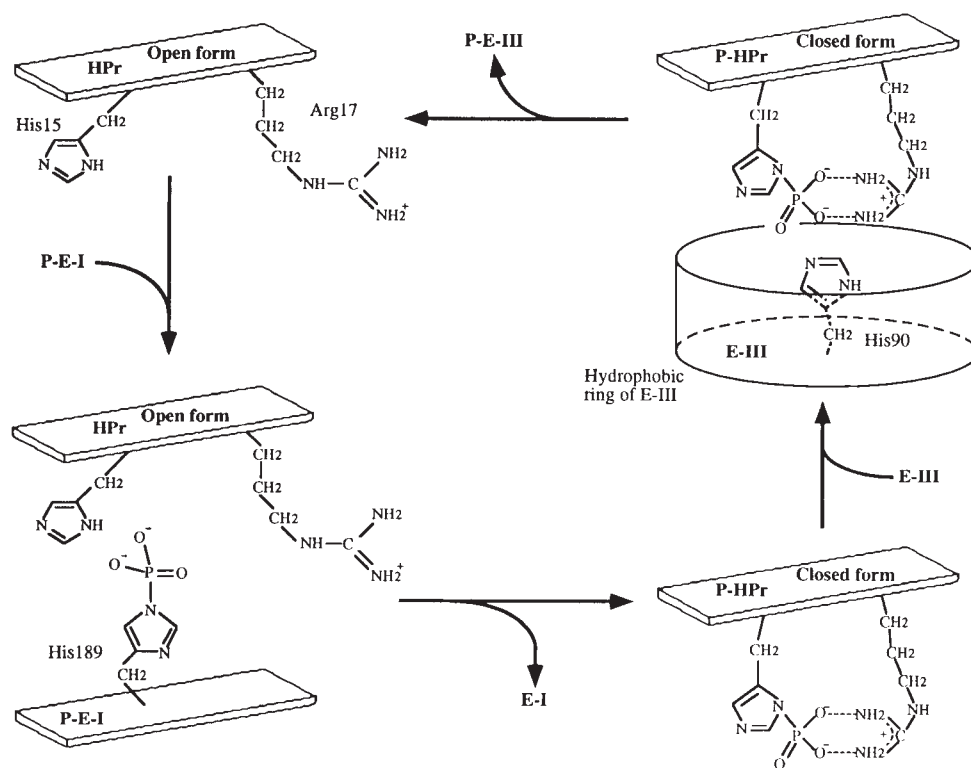


FIG. 3 Proposed HPr phosphotransfer cycle. P-HPr: phospho-HPr; E-I: enzyme I; P-E-I: phospho-enzyme I; E-III: enzyme III; P-E-III: phospho-enzyme III. The closed conformation of the active centre (~ 10 Å in the longest dimension) would be easily accommodated by the hydrophobic doughnut (~ 18 Å diameter¹⁴) of enzyme III whereas the open conformation (~ 18 Å in the longest dimension) could not be accommodated.



reported 2.0 Å X-ray structure of HPr from *B. subtilis*² illustrates the HPr structure with a sulphate ion bound in the active centre, which resembles the closed conformation, namely the phosphorylated state of HPr. The closed conformation in this latter structure does not apparently have torsion-angle strain, as judged from the stereo diagram of the active centre. This recently reported structure thus provides direct evidence supporting our hypothesis that the release of torsion-angle strain takes place during phosphorylation, in proceeding from the open to the closed conformation.

It has been recently shown that enzyme III from both Gram-positive and Gram-negative bacteria has a well defined binding pocket which is very hydrophobic and is doughnut-shaped¹³⁻¹⁵, with the phosphoryl-accepting residue, His 90, located in a

central depression of the hydrophobic ring. The hydrophobic doughnut and His 90 of enzyme III would precisely direct the approach of P-HPr (Fig. 3). As the phosphoryl group leaves HPr, the charge on Arg 17 is no longer neutralized and HPr is repelled by the extreme hydrophobic environment. Simultaneously, the His 15 and Arg 17 side chains of HPr would separate and the active centre would take on the strained open conformation (Fig. 2a), ready for the next cycle; formation of hydrogen bonds to His 15 and Arg 17 would help to stabilize the open conformation and the protein would be in an overall energy minimum. By cycling in this manner (Fig. 3), HPr could accept and donate a phosphoryl group by adopting an open conformation suitable for interacting with enzyme I and a closed conformation suitable for interacting with enzyme III. □

Received 22 June; accepted 9 November 1992.

- Meadow, N. D., Fox, D. K. & Roseman, S. A. *Rev. Biochem.* **59**, 495-542 (1990).
- Herzberg, O. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2499-2503 (1992).
- Wang, B. C. *Meth. Enzy.* **115**, 90-112 (1985).
- Kalbitzer, H. R. et al. *Biochemistry* **21**, 2879-2885 (1982).
- Hol, W. G. J. *Prog. Biophys. molec. Biol.* **45**, 149-195 (1985).
- Åqvist, J., Luecke, H., Quiocho, F. A. & Warshel, A. *Proc. natn. Acad. Sci. U.S.A.* **88**, 2026-2030 (1991).
- Sharma, S. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 4877-4881 (1991).
- Ramakrishnan, C. & Ramachandran, G. N. *Biophys. J.* **5**, 909-933 (1965).
- Jones, T. A. *J. appl. Crystallogr.* **11**, 268-272 (1978).
- Fersht, A. R. et al. *Nature* **314**, 235-238 (1985).
- van Dijk, A. A., de Lange, L. C., Bachovchin, W. W. & Robillard, G. T. *Biochemistry* **29**, 8164-8171 (1990).
- Herzberg, O. & Moul, J. *Proteins: Struct. Funct. Genet.* **11**, 223-229 (1991).
- Liao, D. et al. *Biochemistry* **30**, 9583-9594 (1991).

- Worthylake, D. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10382-10386 (1991).
- Stone, M. J. et al. *Biochemistry* **31**, 4394-4406 (1992).
- Messerschmidt, A. & Pflugrath, J. W. *J. appl. Crystallogr.* **20**, 306-315 (1987).
- Dickerson, R. E., Weinzierl, J. E. & Palmer, R. A. *Acta crystallogr.* **B24**, 997-1003 (1968).
- Evans, S. V. *J. molec. Graphics* (in the press).
- Wittekind, M. G., Reizer, J. & Klevit, R. E. *Biochemistry* **29**, 7191-7200 (1990).
- Klevit, R. E. & Waygood, E. B. *Biochemistry* **25**, 7714-7781 (1986).
- Hammen, P., Waygood, E. B. & Klevit, R. E. *Biochemistry* **30**, 11842-11850 (1991).
- Kalbitzer, H. R., Neidig, K.-P. & Hengstenberg, W. *Biochemistry* **30**, 11186-11192 (1991).
- van Nuland, N. et al. *Eur. J. Biochem.* **203**, 483-491 (1992).
- El-Kabbani, O. A. L., Waygood, E. B. & Delbaere, L. T. J. *J. biol. Chem.* **262**, 12926-12929 (1987).
- Brünger, A. T., Kuriyan, J. & Karplus, M. *Science* **235**, 458-460 (1987).

ACKNOWLEDGEMENTS. We thank W. Hengstenberg for the *S. faecalis* HPr sample, R. Klevit for making the NMR coordinates of *E. coli* and *B. subtilis* HPrs available to us for comparison, E. B. Waygood for discussion and for encouragement, and L. Prasad and A. Matte for suggestions. L.T.J.D. acknowledges the support of the Medical Research Council of Canada.

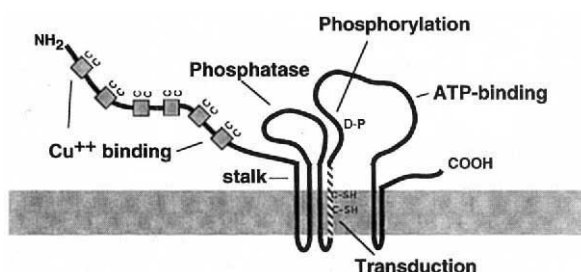
Cloning the Menkes disease gene

Three groups have successfully isolated the gene responsible for the X-linked Menkes disease, heralding great promise for our understanding of copper metabolism and for diagnosis of the disorder.

In 1962, John Menkes, a trainee neurologist at Columbia University in New York (who eight years earlier had described the first case of so-called Maple syrup urine disease) witnessed an infant with an unusual deteriorating brain disease and white, twisted, brittle hair¹. The disease was evidently familial, and those affected lived for only a year or so. Initial suspicions as to the site of the defect focused on amino acid metabolism, but blood and serum measurements provided no clues. Ironically, ceruloplasmin levels monitored in a 1-week-old affected child also proved normal — unknown to Menkes at the time, they would subsequently decline, providing a telling clue to the underlying defect. But it was not until ten years later that David Danks noted that the unusual hair in Menkes disease, as it was then known, resembled the wool of sheep in Australia that suffered from copper deficiency². (Copper is required for the cross-linking of amino acids in keratin and other connective tissue proteins.)

More recent biochemical studies have indicated that Menkes disease results from a copper-transport abnormality. Copper (Cu) ions are vital cofactors for at least half-a-dozen enzymes throughout the body, including lysyl oxidase, cytochrome c oxidase, superoxide dismutase and dopamine β -hydroxylase, and the majority of Menkes disease symptoms can be explained rather satisfactorily by deficiencies in these enzymes. Menkes patients, however, possess elevated amounts of Cu in most tissues (except the liver) suggesting that the intracellular transport of Cu is at fault. In particular, studies with cultured fibroblasts from Menkes patients have

shown the efflux of Cu to be defective³. Speculation can now be firmly put to rest with the accounts in this month's *Nature Genetics* of the cloning of the defective Menkes disease gene on the X chromosome. Like most successes using the positional cloning strategy⁴, the characterization of Menkes patients with translocations (of the long arm of the X chromosome) provided the key to finding the gene. And, fortunately, the complete structure of the Menkes gene product (see figure) contains a bounty of information for those seeking to understand ion transport and the role of heavy



Predicted structure of the Menkes disease protein — a P-type ATPase (see ref. 5).

metals in cellular function.

The gene, which has been independently isolated by the groups of Jane Gitschier and Seymour Packman (University of California, San Francisco), Anthony Monaco (Institute of Molecular Medicine, Oxford) and Thomas Glover (University of Michigan), detects an 8.5 kilobase (kb) messenger RNA in most tissues examined⁵⁻⁷ (although notably only minute levels are found in the liver). But in 23 out of 32 Menkes patients examined in all, the RNA was either qualitatively or quantitatively abnormal⁵⁻⁷. The Oxford group also found that 16 of 100 patients had various non-overlapping deletions of the gene⁶.

The full-length complementary DNA sequence obtained by the San Francisco team reveals a predicted 1,500-amino acid protein containing six transmembrane domains⁵. Database homologies indicate that the Menkes gene product is a member of a cation-transporting P-type ATPase subfamily. The greatest resemblance is to the Cu export protein of *Enterococcus hirae*, but there are significant similarities to other prokaryotic cation transporters including those for

cadmium and potassium. The predicted structure contains typical P-type ATPase ATP-binding, phosphorylation and phosphatase domains, as well as a cation channel. The phosphorylation domain contains a conserved aspartic acid residue which is thought to undergo a round of phosphorylation during the transport cycle. The protein also possesses a conserved proline residue (flanked by conserved cysteine residues) thought to be involved in energy transduction and ion binding. The N terminus contains six 23-residue repeats, each containing a Cys-X-X-Cys motif and probably involved in copper binding.

The cloning of the Menkes gene raises many intriguing questions. Could Wilson's disease, another hereditary disorder of Cu metabolism in which the liver accumulates Cu while other tissues are relatively unaffected, be attributable to a defect in a related liver-specific transporter? How similar is the gene responsible for *mottled*, the murine homologue of Menkes disease? And could the Menkes gene also be responsible for another disorder, X-linked cutis laxa (occipital horn syndrome), a connective tissue disorder featuring bladder abnormalities and bony 'horns' for which Cu abnormalities have been described? The San Francisco group have found that cultured fibroblasts from two cutis laxa patients are virtually lacking the Menkes gene transcript (although Southern analysis of the DNA appears normal)⁸. This suggests that the two disorders are indeed allelic, although interestingly, Menkes disease is also commonly associated with severe reductions in RNA levels. Finally, the cloned gene should greatly facilitate diagnosis of Menkes disease, which currently is performed biochemically in just two centres worldwide⁷. Although a severe disorder, early detection and copper histidine replacement therapy on a daily basis offer a promising form of treatment⁹.

Kevin Davies

Kevin Davies is Editor of Nature Genetics.

Also in this month's *Nature Genetics*: expression studies of the fragile X syndrome gene, *FMR-1*, and a putative point mutation in a fragile X patient; mutations of the neurofibromatosis-1 gene in neuroblastoma; and the problems of defining genes in psychiatric disorders as evidenced by the case of X-linked manic depression.

1. Menkes, J. H. et al. *Pediatrics* **29**, 764-779 (1962).
2. Danks, D. M. et al. *Lancet* **1**, 1100-1102 (1972).
3. Hamer, D. H. *Nature Genet.* **3**, 3-4 (1993).
4. Collins, F. S. *Nature Genet.* **1**, 3-6 (1992).
5. Vulpe, C. et al. *Nature Genet.* **3**, 7-13 (1993).
6. Chelly, J. et al. *Nature Genet.* **3**, 14-19 (1993).
7. Mercer, J. F. B. et al. *Nature Genet.* **3**, 20-25 (1993).
8. Levinson, D. et al. *Nature Genet.* **3**, 6 (1993).
9. Kaler, S. G. & Gahl, W. A. *Am. J. hum. Genet.* **51**(suppl.) A170 (1992).